Journal of
Experimental Psychology

VOL. 32, NO. 1 JANUARY, 1943

TESTS OF THE TWO-FACTOR THEORY OF RETROACTIVE INHIBITION

BY JOHN A. McGEOCH[1] AND BENTON J. UNDERWOOD

State University of Iowa

I. INTRODUCTION

As a result of their recent experiment, Melton and Irwin (**9**) have questioned a theory of retroactive inhibition (*RI*) which attributes all of the decrement to the competition of responses during recall and relearning of the original activity. In their experiment, Melton and Irwin discovered that the function relating *RI* and degree of interpolated learning was not the same as the function relating frequency of overt inter-list intrusions at recall and degree of interpolated learning. Assuming that overt intrusions are indices of competition of responses, these writers argue that their results demand one or more additional factors beyond competition to account for all of the measured *RI*. Consequently, they have proposed a two-factor theory.

This two-factor theory allocates a portion of the total decrement in the usual *RI* operations to the competition of response systems at the time of recall and relearning. The remainder has been attributed, at least tentatively, to an active unlearning or permanent weakening of the original responses as a result of their non-reinforcement during the interpolated learning.

It should be emphasized that this second or unlearning factor is in no sense antithetical toward the more general competition theory which has grown up primarily as a result of a long series of analyses made by McGeoch and his associates (**2**, **3**, **4**, **5**, **6**, **7**, **8**). The ultimate influence of the unlearning process is to change the relative strengths of the original and interpolated response systems at recall. The importance of this postulated second factor is in identifying the

[1] This is the personal research of the late Professor John A. McGeoch. The second of the two experiments reported here was in progress at the time of his death. The junior author expresses his appreciation to Professor Kenneth W. Spence, who has critically read the manuscript.

locus of the process which presumably weakens the original *S–R* connections.

A pointed deduction from the two-factor theory states that the recall of the original activity should be less than the recall of an equally well-learned interpolated activity, the equality of learning being measured in terms of levels of acquisition after a definite number of trials. The reasoning is straightforward. If the competition of responses at the time of attempted recall is the only factor producing inhibition, the decrement should be approximately equal for both original and interpolated activities. On the other hand, if during the interpolated learning a weakening of the original responses occurs (the unlearning factor), the recall of the interpolated activity should be greater than the recall of the original activity. The decrement in the recall of the original activity after interpolation defines *RI*, whereas the decrement in the recall of the interpolated activity after prior learning of another activity defines proactive inhibition (*PI*) (12).

This deduction was first stated and tested by Melton and von Lackum (10). Serial learning of 10 consonant syllables was employed in testing the deduction. Conditions were arranged so that separate measurements of *RI* and *PI* were obtained both with a high degree and low degree of similarity between the original and interpolated lists. With both degrees of similarity the deduction was confirmed, the recall of the interpolated list being greater than the recall of the original list.

The two experimental studies to be reported here are attempts to test the generality of the deduction by using the paired associates technique with two-syllable adjectives making up the pairs. In the first experiment (Experiment *A*), paired associates learning was enforced by varying the order of presentation of the pairs on each trial. The basic design of the Melton and von Lackum (10) experiment was retained. These operations produced differences between *RI* and *PI* which were small and statistically insignificant, though in the direction that would tend to confirm the deduction. Such results led to the belief that the reduction in the differences between *RI* and *PI* in our experiment, as compared to the Melton-von Lackum experiment, was a major function of the method of learning; *i.e.*, whether serially or by pairs. The second experiment (Experiment *B*) was performed with the adjectives presented in the same serial order on each trial in an effort to combine the two forms of learning within the single experiment.

II. Method and Materials

Experiment A.—Table I gives the outline of the four experimental conditions used to measure *RI* and *PI*. Conditions I and II provide the control and experimental conditions, respectively, for determining *RI*, and Conditions III and IV, for determining *PI*.

TABLE I

Outline of Experimental Conditions Used to Measure Retroactive and Proactive Inhibition

Experiment A

Conditions	Original Learning	Rest	Interpolated Learning	Rest	Relearning
I RI (control)	6 trials List A–B	1 min.	——	30 min.	Relearn List A–B
II RI (exp.)	6 trials List A–B	1 min.	6 trials List A–C	24½ min.	Relearn List A–B
III PI (control)	——	—	6 trials List A–C	24½ min.	Relearn List A–C
IV PI (exp.)	6 trials List A–B	1 min.	6 trials List A–C	24½ min.	Relearn List A–C

All lists, whether of the original or interpolated learning, were presented for 6 trials before the major rest period. In Conditions I and II relearning took place after 30 min, and after 24½ min, in Conditions III and IV. Because of this time difference between learning and recall in the experimental conditions (II and IV), we have followed Melton and von Lackum (10) in providing separate control conditions.[2] Relearning under all conditions was to a criterion of 2 successive errorless trials.

In Table I the designation *A–B* under the original learning represents a pair of adjectives, while the *A–C* under the interpolated learning indicates another pair of adjectives in which the stimulus, *A*, of the original learning is paired with a new response, *C*, in the interpolated list. Each list consisted of 10 pairs of two-syllable adjectives. The lists were so constructed that none of the initial letters of the stimulus words in any one list was identical with another, nor was there any identity in the first letters of the response words. The same criterion was applied in eliminating similar syllables. In addition, no identical first letters or similar syllables were used in any *A–B*, *A–C* pairings. Thus, an attempt was made to eliminate all similarity factors except that inherent in the *A–B*, *A–C* relationship.

The pairs were presented at a 4-sec. rate on a modified Hull memory drum. This 4-sec. period allows for a 2-sec. anticipation period in which the stimulus word is exposed alone, and a 2-sec. period in which both the stimulus and response words are shown together. Between each pair of trials 8 sec. elapsed. Rest periods over 1 min. in length were filled by having the subject rank cartoons in the order of their humor.

The *S*s were 24 college undergraduates, 19 women and 5 men. With four experimental conditions a complete counter-balancing design is attained. Four sets of experimental lists were used, the same set being used for each *S* on any specific day regardless of the condition under which he worked. That is, all *S*s used the lists making up Set No. 1 on the first day, Set No. 2 on the second day, etc. Since we have no knowledge of the relative difficulty of the lists used, aside from the equalization resulting from the criteria employed in their construction, this method

[2] It should be mentioned that this time difference is not impossible to avoid. The conditions could be so arranged that the interpolated list is recalled after the *same* lapse of time as the original list. Thus:

A–B, *A–C*..... rest..... *A–B* (recall)
A–B, *A–C*..... rest..... *A–C* (recall)

of counter-balancing conditions while keeping the lists constant assures only that influences of the unknown differences in the lists will be fairly evenly distributed over all conditions. As a further precaution, the original and interpolated lists were reversed for the second group of 12 *Ss*, so that the lists which had formerly been used as the original lists, were then employed as the interpolated lists.

The instructions to the *Ss* stressed the importance of making associations between the pairs by treating the pairs as if they composed a vocabulary list in which the second word was a translation of the first. In addition, the *Ss* were cautioned against learning the pairs in the order in which they occurred on the first trial, being told that the order would be changed on each trial. Actually, 5 different orders were used as follows:

Order No. 1	1	2	3	4	5	6	7	8	9	10
Order No. 2	3	6	9	5	8	1	10	2	4	7
Order No. 3	8	4	2	6	10	7	1	5	3	9
Order No. 4	6	3	7	9	1	8	4	10	2	5
Order No. 5	2	9	6	1	7	5	8	3	10	4

The *Ss* were never told which list they would be expected to recall until just prior to the starting of the drum at the end of the major rest period. The extra list (*A–C* for Condition II, and *A–B* for Condition IV) was always relearned 5 trials after the relearning of the primary list in an effort to keep the motivation constant in the learning of the original and interpolated lists. Equal learning in terms of trials was maintained for each experimental day by requiring the *Ss* to learn an extra list for 11 trials on Conditions I and II (*A–C* for Condition I, and *A–B* for Condition III). This learning occurred after the relearning of the primary list. Two practice sessions preceded the 4 experimental days. Each session, whether practice or experimental, required about 1 hour.

Response latency was measured by a voice-key and Springfield clock. The complete latency data will not be presented here, although certain tendencies in response speed will receive reference.

Experiment B.—As mentioned previously, Experiment *B* varied from Experiment *A* primarily in that the adjectives were presented in the same order on each trial. The same sequence of conditions was used, with one exception: the original and interpolated lists were presented for only 4 trials, rather than 6 trials as in Experiment *A*. This change in the criterion of trials was made to equalize the degree of learning as determined by the frequency of correct anticipations occurring on the last trial of the original and interpolated lists. Thus, with the adjectives presented in different orders (Experiment *A*), the degree of learning by this criterion at the end of 6 trials was the same as the degree of learning at the end of 4 trials when the pairs were presented in a constant order (Experiment *B*). (See Table II.)

The use of only 4 trials for the interpolated list increased the major rest period on Conditions II, III, and IV to slightly over 26 min. The instructions to the *Ss* of Experiment *B* were the same as for Experiment *A*, except that no indication was given as to whether or not the words would be presented in the same serial order on each trial, and they were not instructed to learn the pairs in pairs. In this group of 24 *Ss*, all undergraduates, there were 16 men and 8 women.

III. Results

Learning of the Original and Interpolated Lists.—The degree of learning of the original and interpolated lists was controlled by a trial criterion rather than by an acquisition criterion. Before any inter-condition comparisons can be made, it is necessary to show the level of learning attained on all original and interpolated lists. Table II presents the mean number of correct anticipations for all trials of the original and interpolated lists and for the last trial of these lists. These data will be considered separately for each experiment. It should be remembered that 6 trials made up the original and interpolated learning for Experiment *A*, and only 4 trials were used for Experiment *B*.

TABLE II

MEASURES OF THE DEGREE OF LEARNING OF THE ORIGINAL (*OL*) AND INTERPOLATED (*IL*) LISTS: MEAN NUMBER OF CORRECT ANTICIPATIONS ON THE LAST TRIAL AND MEAN TOTAL CORRECT ANTICIPATIONS ON ALL TRIALS

	Experiment *A*				Experiment *B*			
	Last Trial		All Trials		Last Trial		All Trials	
Condition	*M*	σ_M	M	σ_M	*M*	σ_M	*M*	σ_M
I (OL)	5.54	.49	19.42	1.99	5.50	.52	12.58	1.33
II (OL)	5.58	.50	19.21	-2.01	5.67	.53	12.83	1.33
(IL)	6.21	.49	21.83	2.29	5.46	.35	12.29	.91
III (OL)	5.13	.52	19.33	2.31	5.46	.45	13.25	1.23
IV (OL)	5.75	.46	19.79	2.21	6.33	.49	13.79	1.27
(IL)	5.83	.48	20.08	2.16	5.46	.43	11.62	.94

Experiment A.—As measured by the mean correct anticipations on the last trial (6th), the original learning is nearly equal for all conditions. The largest difference is between Conditions III and IV, being .62 ± .47 (t = 1.32).[3] When the mean total anticipations on all trials is considered, the greatest difference obtains between Conditions II and IV. This difference is .58 ± 1.66 (t = .35).

A direct comparison of *RI* and *PI* necessitates nearly equal degrees of learning for the original list of Condition II and the interpolated list of Condition IV, since these are the lists to be compared at recall and relearning. The mean difference in the frequency of correct anticipations on the last trial of these conditions is .25 ± .46 (t = .54). The difference in the total correct anticipations for all trials is .87 ± 1.90 (t = .46). Thus, the level of acquisition for these lists may be considered equal for later comparisons.

No associative inhibition [4] is evident in the learning of the second list, whether measured by the anticipations on the last trial or by the total anticipations on all trials. In fact, the means suggest that there may be slight facilitation as a result of having learned the first list. However, a comparison of the mean correct anticipations on the last trial of the original learning for Condition II, with the same measure for the interpolated learning of that condition, yields a difference of only .63 ± .45 (t = 1.40). In terms of total correct antici-

[3] All measures of reliability are sigma of the means. Correction has been made for the correlation between the measures. With 23 degrees of freedom, a t value of 2.807 is taken as a very significant mean difference (1).

[4] We are using the older term, associative inhibition, to designate the retardation in the learning of a list as a consequence of having learned a prior list. This phenomenon has also been called proactive inhibition in learning (9), or by the more generic term, negative transfer.

pations, the difference is 2.62 ± 1.54 (t = 1.70). Thus, any associative inhibition which might retard the learning of the second list is offset by positive transfer.

If associative inhibition is present, it should be evident on the first anticipation trial of the interpolated list. Each subject learned 4 original lists and 2 interpolated lists. The mean of the individual means for the first anticipation trial of the original learning is 1.99, and for the corresponding trial of the interpolated list, 1.69. The difference is .30 ± .17 (t = 1.77). Associative inhibition appears to be insignificant in paired associate learning with the degree of original learning used here.

Experiment B.—In terms of the number of correct anticipations on the last trial (4th), the level of acquisition for the original and interpolated lists is remarkably equal except for the aberration which appears in the original learning of Condition IV. The difference between this mean and the mean of Condition III (lowest level of acquisition) is .87 ± .38, a difference which is 2.29 times its sigma. There is no ready explanation for this difference. But this unexpected level of acquisition is not crucial in prohibiting further comparisons. Since it occurs on the original learning of Condition IV (proactive experimental), no measure of recall is taken. Any influence this experimental anomaly might have would be to change the associative inhibition in learning the second list, or to change slightly the competitive strength of the original and interpolated responses at recall. The second point will be considered later; the first, at once.

The difference between the mean correct anticipations on the fourth trial of the original and interpolated lists for Condition IV is .87 ± .61 (t = 1.43). Likewise, there are consistently fewer total correct anticipations on the interpolated learning than on the original learning. Thus by this measure, the difference of the original and interpolated learning for Condition IV is 2.17 ± 1.33 (t = 1.63). However, the degree of learning achieved on the fourth trial of Condition IV is equivalent to the level achieved on the interpolated list of Condition II, and we may conclude that any associative inhibition accruing is exceedingly transitory. This is in keeping with previous findings (**9**).

Associative inhibition as measured on the first anticipation trial has been calculated in the same manner as in Experiment *A*. The mean of the individual means for Experiment *A* is 2.98 for the original learning and 2.19 for the interpolated learning. The difference, .79 ± .35, is 2.26 times its sigma. While not highly significant, this difference suggests greater associative inhibition than was found for Experiment *A*.

Retroactive and Proactive Inhibition at Recall.—Table III shows the mean correct anticipations on the first three relearning trials, together with the percentages of *RI* and *PI* calculated by the formula, 100 (Rest-Work)/Rest. The percentages are plotted in Fig. 1. The results of each experiment will be considered separately.

TABLE III

MEAN CORRECT ANTICIPATIONS ON THE FIRST THREE RELEARNING TRIALS, AND THE PERCENTAGES OF *RI* AND *PI*

Experiments A and B

Experiment *A*

	Trials					
	1		2		3	
Condition	*M*	σ_M	*M*	σ_M	*M*	σ_M
I	4.54	.50	6.25	.53	6.63	.68
II	2.50	.36	4.71	.50	5.71	.48
	% *RI* 44.9		% *RI* 24.7		% *RI* 13.9	
III	4.79	.55	5.92	.55	6.46	.54
IV	3.25	.45	5.33	.49	6.21	.47
	% *PI* 32.2		% *PI* 10.0		% *PI* 3.9	

Experiment *B*

	Trials					
	1		2		3	
Condition	*M*	σ_M	*M*	σ_M	*M*	σ_M
I	3.71	.53	5.58	.56	7.13	.50
II	1.71	.39	4.92	.46	6.46	.44
	% *RI* 53.9		% *RI* 11.8		% *RI* 9.4	
III	3.33	.46	5.67	.40	7.08	.35
IV	2.54	.35	5.92	.40	7.33	.50
	% *PI* 23.7		% *PI* −4.4		% *PI* −3.5	

Experiment A.—*RI* is determined by a comparison of the recall on Condition I (control) and Condition II (experimental). The mean difference on the first relearning (recall) trial is 2.04 ± .39. This difference is 5.23 times its sigma and hence highly significant. On the second relearning trial the difference is still significant, being 1.54 ± .50 (t = 3.08), but *RI* has dissipated by the third trial, the difference being only .92 ± .52 (t = 1.77). The respective percents of *RI* are 44.9 percent, 24.7 percent, and 13.9 percent.

A comparison of Condition III (control) and Condition IV (experimental) determines the amount of *PI*. On the first relearning

trial the difference is 1.54 ± .48 (t = 3.21), and on the second trial, .59 ± .44 (t = 1.34). Thus, while significant *PI* is evident on the first relearning trial, the difference is negligible on the second trial. The percentages of *PI* for the three relearning trials are 32.2 percent, 10.0 percent, and 3.9 percent, respectively.

A direct comparison of the recall under Conditions II (*RI*) and IV (*PI*) requires an initial comparison of the recall on the two control conditions (I and III) because of the slight time difference between learning and recall for these two conditions. From the data of Table III it can be shown that a direct comparison is warranted. On the first recall trial the difference is .25 ± .43 (t = .58) and on the second, .33 ± .43 (t = .77).

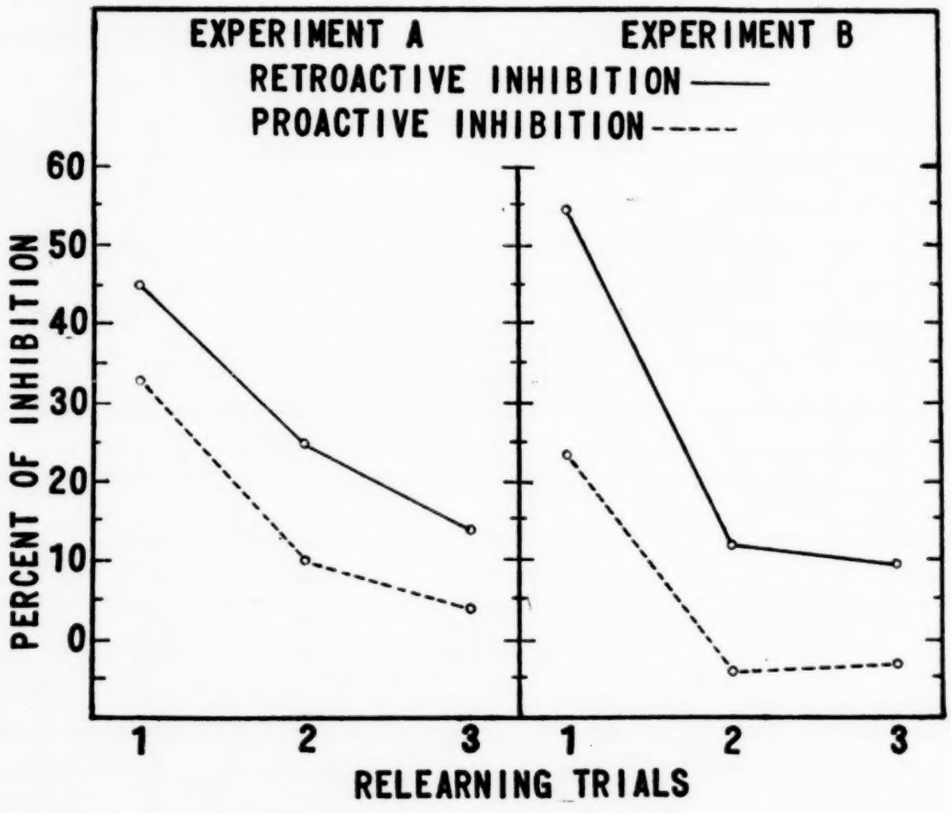


FIG. 1. Percent of retroactive and proactive inhibition on the first three relearning trials.

Fig. 1 would suggest that *RI* is greater than *PI*, thus confirming the deduction from the two-factor theory. The mean difference on the first relearning trial is .75 ± .37 (t = 2.03), and on the second trial, .62 ± .42 (t = 1.48). These differences, while small and statistically insignificant, are consistently in the direction that would support the deduction, and confirm the Melton-von Lackum (**10**) study. The discrepancy between these present findings and those of the above mentioned writers lies in the *reduction* in the difference between *RI* and *PI* in the present experiment as compared with their study.

Experiment B.—The results of Experiment *B* can be stated briefly. Of all the inter-condition comparisons, there is only one significant difference. The difference between Conditions I and II on the first recall trial is 2.00 ± .48 (t = 4.17), indicating significant amounts of *RI* at this locus. *RI* has almost entirely dissipated on the second recall trial, the difference being .66 ± .67 (t = .99). *PI*, while undoubtedly present, is small, the difference on the first trial being only .79 ± .55 (t = 1.44).

The direct comparison of the recall under Condition II and under Condition IV yields a mean difference of only .83 ± .48 (t = 1.73), but again the difference is in the direction that would be predicted by the two-factor theory.

It is apparent from Table III and Fig. 1 that the rate of dissipation of *RI* is faster for Experiment *B* than for Experiment *A*. Also, while the level of acquisition in terms of the mean correct anticipations on the last trial of the original and interpolated learning is virtually the same, the absolute decrement in Experiment *B* is greater than in Experiment *A* for all conditions. This would lead to the generalization that forgetting was less when the adjectives were presented in different orders on each trial than when the pairs were presented in the same order, but that the rate of recovery is less rapid. It is not entirely clear from the present data why this should be. It is, probably, a function of the differences resulting from the serial learning which occurred in Experiment *B*. These factors will be discussed more fully later.

Retroactive and Proactive Inhibition in Trials to Relearn.—Data on the number of trials required to relearn the lists under all conditions to the two criteria of mastery are shown in Table IV. No discussion of these data is necessary. None of the differences between the means are significant, and many of the differences are not in the expected direction. These data tend to bear out previous findings (10) that

TABLE IV

RI AND *PI* AS MEASURED BY TRIALS TO RELEARN TO 1 PERFECT AND 2 SUCCESSIVE PERFECT TRIALS

Condition	Experiment *A*				Experiment *B*			
	1 Perfect		2 Perfect		1 Perfect		2 Perfect	
	M	σ_M	*M*	σ_M	*M*	σ_M	*M*	σ_M
I	10.18	1.61	13.08	2.17	6.91	1.03	7.88	1.16
II	10.63	1.69	13.46	2.04	6.27	.44	7.04	.68
III	8.13	1.12	11.42	1.66	6.38	.69	7.29	.81
IV	8.13	.96	10.38	1.28	5.75	.71	7.13	.90

RI and *PI* as measured by the recall scores provide a much more sensitive measure than do trials to relearn to a criterion of perfection.

Specific Inter-List Intrusions.—Recent studies (**8, 9, 10**) have stressed the importance of overt, inter-list intrusions for the development of a precise theory of *RI* and *PI*. Undeniably, the appearance of a word from the interpolated list during the relearning of the original list (*RI*) and the appearance of a word from the original list during the relearning of the interpolated list (*PI*) constitute unequivocal evidence of the transfer of responses which *may* produce a decrement in retention or retard the learning.

TABLE V

FREQUENCIES AND LOCI OF SPECIFIC OVERT INTER-LIST INTRUSIONS DURING RELEARNING

Condition	Relearning Trial			Remainder	Total
	1	2	3		
Experiment *A*					
RI (Cond. II)	16	3	3	6	28
PI (Cond. IV)	9	4	1	1	15
Experiment *B*					
RI (Cond. II)	13	4	1	2	20
PI (Cond. IV)	5	0	1	0	6

Table V gives the total frequency of intrusions occurring during the relearning, with the specific frequencies for each of the first three relearning trials. These data show rather clearly the correlation which exists between the frequency of intrusions and the amount of inhibition when the original and interpolated activities are learned to a *nearly equal degree*, as they were in these experiments. This is true in spite of the fact that the total number of intrusions which occur are relatively small, as in only one condition (II, Exp. *A*) is the frequency per subject greater than one.

Our recall data supported the generalization that *RI* is consistently greater than *PI*, although the differences are small. If the intrusions may be accepted as indices of inhibition, the generalization is further substantiated, since more intrusions occur under the *RI* conditions than under the *PI* conditions.

Most of the intrusions which occurred in these experiments using paired associates and the *A–B*, *A–C* paradigm, are of the 'paired' type; *i.e.*, *B* is given for *C* at recall (*PI*) or *C* is given for *B* (*RI*). The factors which produce the few intrusions which are not of this type are difficult to determine. Formal similarity of either the stimuli or

of the two responses may be one factor, although such similarities were eliminated as much as possible in the construction of the lists. It is not impossible that some serial position factors may be involved in Experiment *A* in spite of the fact that five different orders would seem to preclude such associations. In Experiment *B* all non-paired intrusions (4 for *RI* and 1 for *PI*) occurred at a position only one serial step away from the locus that would have made them paired intrusions. Thus, serial position similarity seems to play at least a minor role when the pairs are presented in a constant order.

It is apparent from the data presented that all subjects do not show intrusions. In Experiment *A* the total of 28 intrusions for the *RI* condition was divided among 17 subjects, and of the 15 for the *PI* condition, among 7 subjects. In Experiment *B*, 13 subjects gave the total of 20 intrusions for the *RI* condition, and 6 subjects gave the 6 intrusions on the *PI* condition.

Intrusions from the first list also occurred during the *learning* of the second or interpolated list. If these intrusions are indicative of inhibition, we should expect few if any such intrusions in the present experiment, since little evidence was found for associative inhibition. During Experiment *A* a total of 11 such intrusions was observed for Conditions II and IV combined, while in Experiment *B*, 7 was the combined total. The consideration that these intrusions are scattered over 5 trials for Experiment *A* and 3 for Experiment *B* minimizes their importance. Any inhibition which these intrusions represent is exceedingly transient.

IV. Discussion

The results of the present experiments are, as has been seen, in agreement with the implication of the two-factor theory that *RI* will be greater than *PI*. But whereas Melton and von Lackum (**10**) found this difference to be significant in the case of serial learning, the differences were statistically insignificant in both of the present situations.[5] A comparison of the amounts of difference in *RI* and *PI* in all three situations can be most conveniently made in terms of a measure which is obtained by calculating the ratio between the absolute recall scores: 100 (Recall *RI*/Recall *PI*). Melton and von Lackum report the following figures from their experiment: for similar lists, 45.1 percent, 67.9 percent and 78.9 percent for relearning trials 1 through 3, respectively, and for dissimilar lists, 60.3 percent, 95.7 percent and 95.4 percent. The smaller the percentages the greater is the disparity between *RI* and *PI*. In the present experiments, the comparable

[5] Complete confirmation of the results of Experiment *A* were obtained in a duplicate experiment performed by the students of the graduate experimental course in human learning at the University of Iowa.

figures are: for Experiment *A*, 76.9 percent, 88.4 percent, and 91.9 percent; for Experiment *B*, 67.3 percent, 83.1 percent and 88.1 percent.

Although the degree of original and interpolated learning was not the same in the present experiments as it was in the Melton-von Lackum experiment, the comparison of the relative amounts of *RI* and *PI* may be made with some confidence. Such comparisons leave little doubt that on the first recall trial at least, the difference in *PI* and *RI* in Experiment *A* is less than the comparable differences in either of the two Melton-von Lackum experiments using serial learning. These data, and the evidence from Experiment *B* in which the difference between *RI* and *PI* was increased, would suggest that the critical factor related to the amount of difference between *RI* and *PI* is whether the material is learned serially or in pairs.

The suggestion was made earlier that serial learning was involved to some extent in Experiment *B*, which, if true, would be further support for the view that the serial form of learning is the critical variable. There are several reasons supporting the interpretation that serial learning was employed by many of the subjects, either in learning the entire list or in learning portions of the list in this experiment. (1) Many of the subjects reported that they learned the lists serially. (2) The lists were learned faster in Experiment *B* as compared with Experiment *A*. This is in agreement with the findings of McGeoch and McKinney (**5**). (3) More anticipatory errors occurred in Experiment *B* than in Experiment *A*. These errors were usually one or two serial positions removed from the correct position. (4) The initial and final portions of the lists were learned first, a phenomenon which is characteristic of serial learning. (5) Last, the mean latency of response was longer in Experiment *A* than in Experiment *B*. In Experiment *A* the mean latency on all four conditions on the last trial of original learning was 1.43 sec, whereas in Experiment *B*, at the same locus, the mean latency was 1.22 sec. In paired associates learning the subject cannot anticipate a response before the stimulus word for that response is presented. In serial learning, on the other hand, the subject may 'think ahead' for several items.

If these facts are accepted as sufficient evidence for at least partial serial learning in Experiment *B*, and if the previous comparisons between the Melton-von Lackum experiment and Experiment *A* are justified, there is at once established a positive relationship between the extent to which serial learning is involved and the divergence in *RI* and *PI*. Since a few of the subjects in Experiment *B* did not show evidence of serial learning, it would be predicted that had the instructions specifically called for serial learning, the disparity between *RI* and *PI* would have been still larger.

There are certain implications instituted by this inferential relation between the divergence in *RI* and *PI* and serial learning. If the responses in Experiment *B* were in part learned serially, the importance of the similarity factor of the *A–B*, *A–C* paradigm would be reduced, and the factor of serial position similarity would be introduced. The empirical results of Experiment *B* indicate that this change produces a reduction in *PI* and makes *RI* more transitory, although the latter form of inhibition is still as potent on the first recall trial as when the adjectives are learned as paired associates. It would seem to follow that the competition of responses due to serial position factors is more transitory than inhibition attributable to the similarity inherent in the *A–B*, *A–C* relationship.

The transitoriness of the interference allotted to serial position factors receives confirmation from two experimental sources: (1) overt inter-list intrusions that could be identified as due to serial position similarity appear extremely early in the relearning process as compared with those based on meaningful similarity (**9**, **10**); (2) with the original and interpolated lists very dissimilar, *RI* is very transitory (**10**). If serial lists of nonsense syllables are dissimilar, all factors leading to interference are minimized, except that of serial position similarity.

All of the above considerations have a definite bearing on the two-factor theory of *RI*. The present observations have indicated that the second, or unlearning factor, is a legitimate construct, one which would account for the differential strengths of the original and interpolated response systems at the time of recall. Furthermore, the present data suggest certain further functional properties of this construct. The specific deduction which the present experiments were designed to test was that *RI* would be greater than *PI*—the recall of the interpolated list would be greater than the recall of the original list. This deduction suggests that a measure of the amount of unlearning or weakening of the original responses is the difference obtaining between *RI* and *PI*. Since this difference was found to be reduced with paired associates learning as compared with serial learning, the inference is that unlearning is less with paired associates than with serial learning.

However, this inference that unlearning is greater with serial lists than with paired associates is at present purely a conjecture, since there are factors which equivocate such a precise interpretation. In the first place, the intentional confounding of serial and paired associates learning in Experiment *B* prohibits a clear-cut analysis. Thus, a single subject may have learned some words serially and some by pairs, or different subjects may have used each method exclusively. Experiment *B* does seem to have produced the general results in-

tended—that of providing data on material not entirely learned serially and not entirely learned as paired associates. Thus the mode of learning used by the subjects stands somewhere between strict paired associates learning as in Experiment *A* and strict serial learning as in the Melton-von Lackum experiment. Under such conditions, an unambiguous statement of the interfering factors which operated in Experiment *B* is not possible.

Secondly, the fact that *PI* was reduced in Experiment *B* (as compared with Experiment *A*), while *RI* remained relatively constant on the first recall trial, might be accounted for in at least three different ways. (1) These results might indicate greater unlearning in Experiment *B*. If it is assumed that the amount of *PI* is indicative of the competition between the original and interpolated response systems, then the factors producing this competition are less potent in Experiment *B* than in Experiment *A*. This follows from the empirical fact that *PI* is reduced in Experiment *B* as compared with Experiment *A*. As discussed previously, this reduction in competition seems to be a function of the fact that serial learning was involved to a large extent in Experiment *B*, and that interference or competition due alone to serial position identity is less than the competition accruing when two responses are associated with the same stimulus (Experiment *A*). If unlearning was constant for the two experiments, a reduction in *RI* should be expected in Experiment *B* as compared with Experiment *A*. However, *RI* on the first recall trial is about the same for both experiments. The inference is that if unlearning is measured by the difference in *RI* and *PI*, unlearning must be greater in Experiment *B* than in Experiment *A*. In short, if *RI* is some function of competition and unlearning, and if the competition is reduced while *RI* remains the same, there must be an increase in the amount of unlearning.

(2) It might be postulated that *PI* is less influenced by interference due to serial position identity than is *RI*. If such a postulate was accepted, a reduction in *PI* could be expected in Experiment *B* without the concomitant reduction in *RI*. There, however, is no other sound basis for such a postulation.

(3) It is possible that *PI* is more dependent upon the rate of the presentation of the words than is *RI*. There is some evidence that *PI* is considerably reduced with slower rates of presentation. Thus, serial adjectives presented at a 2.3 sec rate gave no significant *PI*, whereas a 2.0 sec rate did give significant *PI* (11). *RI* has not been shown to vary so sharply within such a limited time range of presentation. In Experiment *B*, if the adjectives were learned serially, the anticipation period is extended to 4.0 sec, and *PI* might be reduced while *RI* remains constant when compared with the constant 2.0 sec anticipation time of Experiment *A*.

These factors indicate some of the complications which must condition any statement that unlearning is greater with serial learning than with paired associates learning. The data do, however, point in the direction of such an interpretation. The additional hypotheses necessary for an elaboration of this interpretation preclude any attempt at a systematic presentation now. Further experimentation will probably clarify some of the complicating factors mentioned above.

In conclusion, it should be re-emphasized that the two-factor theory of *RI* is a transfer theory, and the second factor is primarily an addition to a strict competition-at-recall theory. The postulation of the unlearning factor emphasizes the possibility that a process of weakening of the original responses occurs, thus changing the relative strengths of the original and interpolated response systems.

V. Summary

Two experiments were performed to test the deduction from the two-factor theory of retroactive inhibition which states that proactive inhibition will be less than retroactive inhibition. In the first experiment (Experiment *A*) 24 subjects learned lists of paired two-syllable adjectives presented in 5 different orders to enforce paired associates learning. The interpolated learning required the subjects to associate new responses to the stimulus words used in the original learning. Original and interpolated learning was for 6 trials, with a basic rest period of 30 min between original learning and recall. The second experiment (Experiment *B*) varied from Experiment *A* in that the pairs of adjectives were presented in a constant order on each trial, and the original and interpolated learning was for only 4 trials. The results may be summarized as follows:

1. Significant retroactive inhibition at recall was found in both experiments. Significant proactive inhibition was found at recall in Experiment *A*, but not in Experiment *B*.

2. In both experiments retroactive inhibition was consistently greater than proactive inhibition as measured by recall scores, although the differences were not statistically significant.

3. No inhibition of either type was evident in the mean trials required to relearn the lists to 1 perfect and 2 successive perfect trials.

4. The frequency of overt inter-list intrusions appearing at recall was greater for the retroactive conditions in both experiments. This tends to substantiate the consistently greater inhibition present when the original list is recalled than when the interpolated list is recalled.

5. Associative inhibition during the learning of the interpolated list was not significant in either experiment.

Although the differences between retroactive and proactive inhibition were not significant in either experiment, the consistency of the differences supports the deduction from the two-factor theory of retroactive inhibition. The present data, when compared with data from the similar experiment by Melton and von Lackum, suggest that a critical factor determining the difference between retroactive and proactive inhibition is the degree to which the material involves serial learning: the greater the degree of serial learning, the greater the difference between retroactive and proactive inhibition. Possible factors causing this function have been discussed.

(Manuscript received August 3, 1942)

References

1. Fisher, R. A. *Statistical methods for research workers.* (4th ed.) Edinburgh: Oliver and Boyd, 1932. P. 151.
2. McGeoch, J. A. Studies in retroactive inhibition. VII. Retroactive inhibition as a function of the length and frequency of presentation of the interpolated lists. *J. exp. Psychol.*, 1936, **19**, 674–693.
3. McGeoch, J. A., & McGeoch, G. O. Studies in retroactive inhibition. VI. The influence of the relative serial positions of interpolated synonyms. *J. exp. Psychol.*, 1936, **19**, 1–23.
4. McGeoch, J. A., & McGeoch, G. O. Studies in retroactive inhibition. X. The influence of similarity of meaning between lists of paired associates. *J. exp. Psychol.*, 1937, **21**, 320–329.
5. McGeoch, J. A., & McKinney, F. Studies in retroactive inhibition. VIII. The influence of the relative order of presentation of original and interpolated paired associates. *J. exp. Psychol.*, 1937, **20**, 60–83.
6. McGeoch, J. A., McKinney, F., & Peters, H. N. Studies in retroactive inhibition. IX. Retroactive inhibition, reproductive inhibition and reminiscence. *J. exp. Psychol.*, 1937, **20**, 131–143.
7. McGeoch, J. A., & Sisson, E. D. Studies in retroactive inhibition. XI. The influence of the relative serial positions of interpolated synonyms in twenty-item lists. *J. exp. Psychol.*, 1938, **22**, 547–554.
8. McKinney, F., & McGeoch, J. A. The character and extent of transfer in retroactive inhibition: disparate serial lists. *Amer. J. Psychol.*, 1935, **47**, 409–423.
9. Melton, A. W., & McQueen-Irwin, J. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173–203.
10. Melton, A. W., & von Lackum, W. J. Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
11. Underwood, B. J. The effects of punishment in serial verbal learning. *Proc. Ia. Acad. Sci.*, 1941, **48**, 349–352.
12. Whitely, P. L., & Blankenship, A. B. The influence of certain conditions prior to learning upon subsequent recall. *J. exp. Psychol.*, 1936, **19**, 496–504.

STUDIES IN AUDITORY THEORY

I. Binaural Interaction and the Perception of Pitch [1]

BY WILLARD R. THURLOW

Princeton University

Introduction

It is well known that the psychological variable, pitch, is a function not only of frequency and sound composition, but also of intensity.[2] This dependence of pitch upon intensity is of especial importance in determining the adequacy of any theory of hearing. Thus the theory of hearing which postulates that pitch is a function of frequency of nerve discharge is considered disproved by the fact that the frequency of synchronized nerve discharge depends simply upon stimulus frequency and does not vary with the intensity of the stimulating tone (**19**, **21**).

Recent investigators (**11**, **20**, **21**), extending the work of Zurmühl (**27**), have shown that the pitch change resulting from an increase in intensity is downwards for tones of low frequency, and upwards for tones of high frequency, and that in a region between there is no pitch change. They have attempted an explanation of their results [3] in terms of a place theory of hearing, according to which a pitch change is mediated by a shift in the position of maximal stimulation on the basilar membrane. For example, Fletcher has postulated that the portions of the ear which respond maximally to low frequencies act in a manner similar to stretched strings, whose resonant frequencies increase with increased intensity, and that the portions which respond maximally to high frequencies act in a manner similar to pendulums and tuning forks, whose resonant frequencies decrease with increased intensity. Through this action a low frequency at high intensity stimulates a portion of the basilar membrane which would ordinarily respond to a lower frequency, and a lower pitch results. A high frequency at high intensity stimulates a portion of the basilar membrane which would ordinarily respond to a higher frequency, and a higher pitch results. Since there is no change of pitch with increase of intensity in an intermediate frequency region,

[1] From the Psychological Laboratory of Princeton University. The author would like to express his appreciation to his fellow graduate students for serving as subjects in the experiments.

[2] A concise discussion is given by Fletcher (**12**).

[3] The theories of earlier investigators of the pitch-intensity function (**1**–**9**) will not be discussed; their theories were constructed to explain only a lowering of the pitch with increased intensity.

presumably the part of the basilar membrane responding to these frequencies behaves in such a manner that there is no change in resonant frequency with increased intensity. That one part of the inner ear should act in a manner similar to stretched strings, another part in a manner similar to pendulums and tuning forks, and still another like neither, is, as Fletcher himself points out, still a puzzling problem.

Stevens has suggested that with increasing intensity, the portion of basilar membrane at the position of maximal stimulation undergoes overloading; the excitation on either side then grows at a faster rate. At the same time, a process of skewing of the disturbance takes place, imposed by the overall properties of the ear, such that only one maximum occurs. This skewing causes a shifting of the disturbance away from the region of maximal sensitivity of the ear. Thus a change in pitch occurs which is in a downward direction for low frequencies, and in an upward direction for high frequencies. Steven's theory does not indicate what could cause a skewing of the disturbance on the basilar membrane away from the region of maximal sensitivity of the ear.

The difficulty of explaining the mechanism of the shift in the position of maximal stimulation on the basilar membrane, together with the fact that there is no evidence that such a shift actually takes place, leads one to question the validity of this explanation, and to advance the hypothesis that a change in pitch may take place in the absence of a shift in the position of maximal stimulation. A possible means of testing this hypothesis emerges from a consideration of the nature of binaural as contrasted with monaural hearing. The effect of adding to a tone of given frequency and intensity in one ear a tone of equal frequency and intensity in the other ear should be analogous, within limits, to the effect of increasing the intensity of a given monaural tone. That is to say, a change in the total quantity, density, and pattern of nerve discharge should take place. To the extent that the effect is comparable to that caused by an increase of intensity, we should expect to find the binaural tone different in pitch than the original monaural tone. Furthermore, it can be assumed that when the binaural tone is sounding, the positions of maximal stimulation on the basilar membrane in each ear are the same as when the binaural component in each ear is sounded separately. This assumption can be checked experimentally by measuring the effects of the sound conducted from one ear to the other. It can also be assumed that there is no shift centrally [4] in the position of maximal

[4] The postulation of the mediation of pitch by the central position of maximal stimulation is an extension of the theory of the mediation of pitch by position of maximal stimulation on the basilar membrane. The present writer has made this extension since it seemed to him that implications of such an extension should be considered in a discussion of binaural interaction.

stimulation when there is added to a tone of given frequency and intensity in one ear a tone of equal frequency and intensity in the other ear. There is no *a priori* reason for doubting this assumption. The experimental results themselves, however, may be expected to give an indication of its validity. We thus have a situation in which we should expect to find a change of pitch, yet in which it can be assumed there is no change in the position of maximal stimulation, peripherally or centrally.

An experimental investigation of the effects of binaural interaction upon pitch perception was therefore undertaken, since it appeared that such an investigation would furnish crucial evidence concerning the validity of an explanation of the mediation of pitch by the position of maximal stimulation on the basilar membrane.

Experiment I

With the apparatus described below, experiments were performed to determine the difference in pitch between a monaural tone of given frequency and intensity, and a binaural tone whose monaural components were equal in frequency and intensity to the monaural tone. Since preliminary experiments at 200 cycles revealed a somewhat large variability from one session to another, it was decided to determine the values at a series of sessions, separated by at least a day, in order to find the range within which variations occurred over a period of time. This procedure was carried out at 200 and 400 cycles.

Apparatus

The apparatus is represented in semidiagrammatic form in Fig. 1. The apparatus consisted essentially of an oscillator, an attenuator, and a circuit by which the given tone could be presented

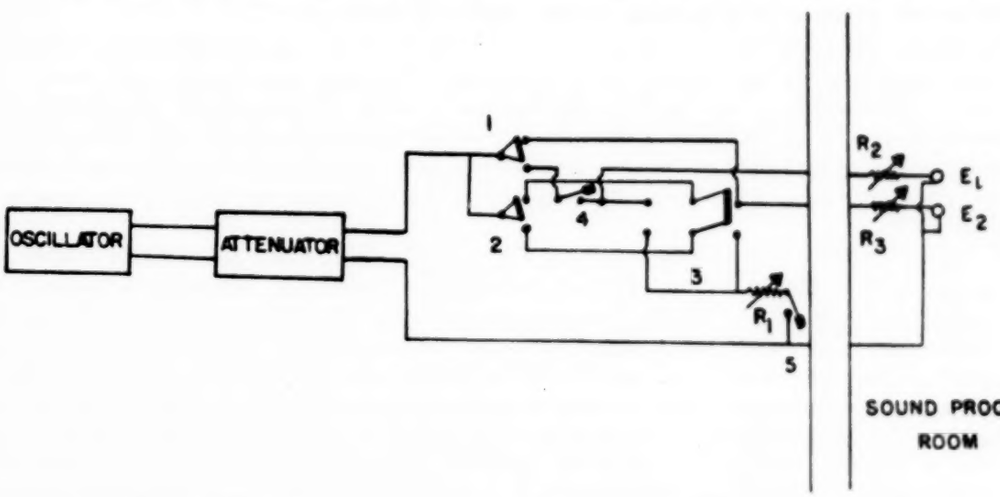


Fig. 1. Representation of the apparatus in semi-diagrammatic form. Numbers 1–5 represent switches; R_1, R_2, and R_3, variable resistors; and E_1 and E_2 the earphones.

monaurally and binaurally in succession. The oscillator had a large dial for crude control of frequency, and also a small dial, calibrated in single cycles, for smaller frequency adjustments. In the diagram, numbers 1–5 represent switches; R_1, R_2, and R_3, variable resistances; and E_1 and E_2 the earphones. Suppose, for the moment, that switches 4 and 5 are closed, and that switch 3 is closed to the right or left; it is evident from the diagram that closing switch 1 will

produce a binaural tone; while closing switch 2 (1 remaining open) will produce a monaural tone in E_1, or E_2, depending upon the direction in which switch 3 is closed. Switches 1 and 2 were ordinary push-button switches which could be operated by means of a simple pressure of the foot.

The procedure for adjusting the apparatus, employed when tests were to be made at a given frequency and loudness level, was as follows: First the earphones were equated in loudness at the frequency and loudness desired. To accomplish this, switches 4 and 5 were left open, and switch 3 was closed to the left; then on closing switches 1 and 2 successively (opening 1 again before 2 was closed) tones were produced successively in the earphones E_1 and E_2 respectively. The tones were equated by adjusting the resistances R_2 and R_3 while the tones were sounded close to one ear. The voltage applied across each phone was then measured. Switch 4 was then closed; and while switch 1 was held closed (thus putting E_1 and E_2 in parallel), the same voltages were produced across the earphones as previously, by adjustment of the attenuator and resistances R_2 and R_3. Secondly, the loudness in the phone which was to be used to produce the monaural tone was made equal to that of the same phone under the binaural condition by making the voltage across the phone equal under the two conditions. This was easily accomplished by closing switch 5, closing switch 3 either to the right or the left, depending on which phone was to be used, and then closing switch 2 while the resistance R_1 was adjusted to give the proper voltage. Voltages were checked at the beginning of each session. The output of the oscillator was such that no amplification was needed.

Procedure

The procedure was essentially one of determining the number of cycles by which the monaural tone had to be changed in order to sound the same in pitch as the binaural tone. A method of constant stimuli was employed; that is to say, the subject was presented with a series of paired tones, the two tones of each pair being presented successively. One of the tones of each pair was a constant standard binaural stimulus, of constant frequency; the other tone was one of a series of variable monaural stimuli. It was the task of the subject to judge whether the second of the two tones of each pair which he heard was higher or lower in pitch than the first tone.

The following special precautions were taken:

a. A given experimental session consisted of two separate series of determinations, the right ear receiving the monaural tone in one series, the left ear receiving the monaural tone in the other series. It was considered necessary to do this to avoid erroneous conclusions which might be drawn because of the possible presence in the subjects of diplacusis.[5]

Switching the monaural tone from one ear to the other was accomplished by switch 3, as described under the section dealing with apparatus. Voltage adjustments were made when necessary. The switching was done in this fashion in order to prevent the possibility of any differences in quality or loudness of the phones causing the results which were obtained. The phones were reversed on the head every other session. By using these controls it was possible to conclude that any consistent difference in the results obtained from the two ears was due to the ears themselves.

b. To avoid fatigue, the tones were sounded for only short intervals of time. Thus, each tone was sounded for .5 of a second, with a .5 second interval between. By using a Kodak Timer, the experimenter was able to present the stimuli with the desired time relations. Preliminary experimentation showed that .5 of a second gave sufficient time for an accurate judgment of pitch to be made. The number of judgments at a given session was kept below 100, in order that the subject's attention might not wander, or actual fatigue set in. This conclusion was reached from preliminary experimentation, on the basis of an examination of the variability of the results during the course of a session, and on the basis of the reports of the observers themselves. During subsequent experiments, a method of tabulation was used such that the values obtained during the first half of the judgments (with the monaural tone in a given ear) could be compared with those obtained during the latter half. Thus any changes due to fatigue or lapse of attention could be detected.

c. The tones used were at a loudness level of 95 decibels.[6] The intensity necessary to produce

[5] Diplacusis is the fairly common condition in which a tone of given frequency and intensity produces a slightly different pitch sensation when sounded successively in the two ears.

[6] Adjustments of sensation level were made to give the 95 decibel loudness level, according to the curves of Fletcher and Munson (13) relating sensation level to loudness level. One adjust-

this loudness was well below the overload point of the phones used. At this intensity, also, the onset of the tone did not produce a disturbing click. From preliminary exploratory experiments it was found that the changes in pitch obtained were the same with long-continued tones. Rubber ear cushions were used on the earphones to minimize the effects of bone and air conduction from one ear to the other.

d. It has been found that the second of two successively sounded tones of the same frequency may sound different in pitch (the so-called "time error"). In the present experiment, therefore, the standard was sounded first in half of the trials. In the other half the variable stimulus was sounded first. Thus an examination of the results could show whether such an error was present.

The actual procedure used by the experimenter in presenting the tones was the following: At the beginning of each of the two series in a given session, the binaural tone and monaural tone were sounded successively, and the monaural tone was adjusted to a frequency which gave a pitch approximately equal to that of the binaural tone. When an approximate value of equality had been obtained, various frequencies about this value from 2 to 4 cycles apart [7] were chosen to provide the various values for the monaural stimulus. In the experiments at a frequency of 200 cycles, seven values were used, and six judgments were taken at each value. For the experiments at 400 cycles, 6 rather than 7 values were used, and eight judgments were made at each value. The stimuli were presented in random order.

Results

a. Results at 200 Cycles

In Table 1 are presented the results of the five subjects. Values of the median change in frequency of the monaural tone necessary to make it equal in pitch to the binaural tone are given for each subject, for each of four experimental sessions. The median value was taken as representing the point of subjective equality; it represents a change in the frequency of the monaural tone such that the monaural tone is judged to be lower in pitch than the binaural tone 50 percent of the time. The values of the median were computed from the data by linear interpolation.

It is evident from the table that the values are negative; this result means that it was necessary to decrease the frequency of the monaural tone in order to make it equal in pitch to the binaural tone; or in other words the binaural tone sounded lower in pitch than the monaural tone. The value of the semi-interquartile range corresponding to each median value is also given in Table 1. As can be seen from these values, the variability within a given experimental session was not large. It can furthermore be seen from the table that the values obtained when the monaural tone was introduced in the right ear are similar to those obtained when the monaural tone was introduced in the left ear. The slight differences which appear in the results for the two ears are probably due to the presence of a small amount of diplacusis.

The average of each 'Right' and 'Left' value for an experimental session was taken, expressed as percent change in frequency, and

ment was used for all subjects; thus the level differed from 95 decibels to the extent that there were slight threshold variations among the subjects.

[7] The value used depended upon the pitch discrimination limen of the subject.

TABLE 1

RESULTS AT 200 CYCLES FOR EACH SUBJECT

The median change in pitch in cycles for right and left ears is shown, together with the semi-interquartile range (Q), for each of four experimental sessions.

Session	Subject *A*				Subject *B*			
	Right		Left		Right		Left	
	Median	Q	Median	Q	Median	Q	Median	Q
1	−10.02	1.95	−19.23	1.86	−11.00	2.96	−12.00	1.23
2	− 8.00	2.04	−13.02	1.63	− 6.00	2.35	− 8.50	0.88
3	− 9.00	3.06	−13.50	6.01	− 7.00	2.23	− 6.00	1.25
4	−10.50	1.14	−16.50	3.00	− 5.00	0.76	− 5.50	2.22

Session	Subject *C*				Subject *D*			
	Right		Left		Right		Left	
	Median	Q	Median	Q	Median	Q	Median	Q
1	−13.50	2.59	−9.00	4.50	−1.80	0.90	−1.50	3.27
2	− 8.00	2.86	−1.50	3.00	0.00	2.65	0.00	2.62
3	− 9.50	1.14	−9.30	1.09	−3.75	1.63	−3.00	2.94
4	−10.00	1.63	−7.80	0.91	−4.20	1.86	0.00	3.48

Subject *E*

Session	Right		Left	
	Median	Q	Median	Q
1	− 4.00	3.92	−6.00	4.50
2	− 4.00	5.58	−5.30	2.60
3	− 5.00	5.80	−7.00	3.50
4	−10.00	2.50	−6.00	1.52

plotted in Fig. 2 for each subject, as a function of the number of the experimental session, in order to show the consistency of each individual from session to session, as well as the differences between individuals. It appears from the graph that the variations within an individual from session to session are smaller than the range of variation between individuals. By use of analysis of variance this can be shown more exactly. Thus, the variability in the shift in pitch when the monaural tone is sounding in the right ear is significantly less within a given subject than between subjects; this result is also obtained when the monaural tone is sounding in the left ear.

From a consideration of these results it can be concluded that a tone of 200 cycles, sounded binaurally at a loudness level of 95 decibels,

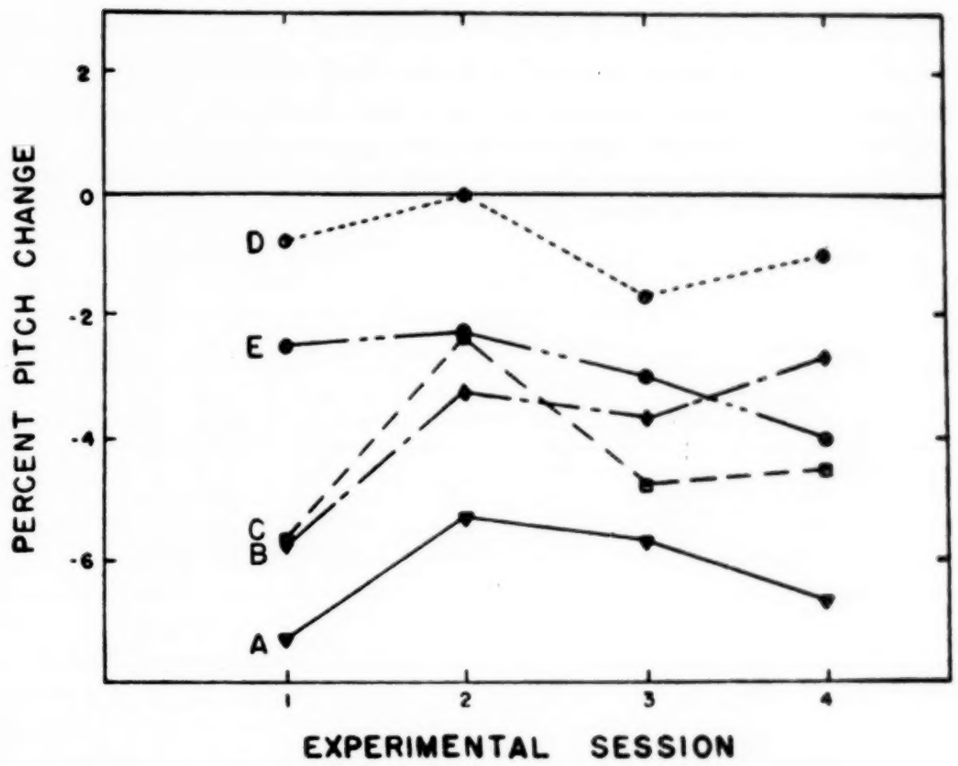


FIG. 2. Results at 200 cycles. The percent pitch change is shown for each subject at each of the experimental sessions.

is perceived consistently as lower in pitch than a monaural tone which is equal in frequency and intensity to each of the binaural components.[8]

b. Results at 400 cycles

The results at 400 cycles are presented in Table 2 and Fig. 3 in a manner similar to that in which the results at 200 cycles were presented, the only difference being that at 400 cycles, 6 experimental sessions were used instead of four.[9] The median values are negative, which fact indicates that, as at 200 cycles, the binaural tone sounded lower in pitch than the monaural tone. After sessions one and two, an interval of more than a week elapsed; then the other four sessions were completed. This procedure was followed to determine whether any large variation in the function occurred over a longer period of time. A consideration of the results shows that no large variation took place. At 400 cycles, as at 200 cycles, the variation shown by a subject at a given experimental session, as measured by the semi-interquartile range, was not large; and the variation for a given sub-

[8] A small number of exploratory experiments performed at 200 cycles at lower loudness levels showed that the change in pitch of three of the subjects tended to decrease, and then to become slightly positive when the loudness level was decreased; while with one of the subjects, the change merely tended to decrease to zero. These results are merely suggestive, since no attempt was made to explore in detail the effect of variation of intensity upon this function.

[9] Subject *E* was no longer available for experimentation.

TABLE 2

RESULTS AT 400 CYCLES FOR EACH SUBJECT

The values given represent the median change in pitch in cycles for right and left ears, together with the semi-interquartile range, for each of six experimental sessions.

Session	Subject *A*				Subject *B*			
	Right		Left		Right		Left	
	Median	*Q*	Median	*Q*	Median	*Q*	Median	*Q*
1	−6.00	2.50	−4.20	1.20	− 7.50	3.50	−7.80	1.20
2	−4.71	0.86	−4.00	1.75	−10.50	3.00	−6.75	4.12
3	−7.50	1.50	−4.80	1.20	− 6.75	2.62	−4.50	3.37
4	−6.00	2.25	−3.00	1.75	− 6.75	1.50	−5.25	1.87
5	−5.00	1.00	−4.00	2.00	− 8.00	3.00	−6.75	1.87
6	−7.50	1.00	−2.25	1.40	− 7.50	3.75	−6.75	3.50

Session	Subject *C*				Subject *D*			
	Right		Left		Right		Left	
	Median	*Q*	Median	*Q*	Median	*Q*	Median	*Q*
1	−4.50	1.10	−3.20	0.80	− 7.00	2.25	−6.00	2.25
2	0.00	1.50	−3.00	2.70	− 3.00	5.00	−7.00	5.00
3	−4.20	1.20	−2.00	1.00	−12.00	3.75	−1.71	0.86
4	−2.25	1.87	−0.66	1.83	−12.00	4.50	−3.60	1.65
5	−6.85	0.57	−3.33	1.83	− 3.00	3.50	−1.50	3.00
6	−2.66	1.33	−4.50	1.25	− 4.50	2.75	−3.00	2.50

ject from session to session was not so great as the variation between subjects (this is shown in Fig. 2). By use of analysis of variance it is possible to demonstrate, as at 200 cycles, that the variation between subjects is significantly greater than the variation within a given subject, for the monaural tone sounding in a given ear. It is possible to conclude, then, that at 400 cycles the subjects consistently judged the binaural tone to be lower in pitch than the monaural tone.

The average pitch change at 400 cycles for all subjects was determined separately for the case when the monaural tone was sounding in the right ear and when it was sounding in the left ear. This was done because of the fact that the two sets of values were dependent, yet not part of the same distribution. These values, which represent the average of the average results for each observer, are 1.52 percent and 1.01 percent, for the monaural tone sounding in the right and the left ears respectively. The corresponding values at 200 cycles are 3.51 percent and 3.51 percent respectively. If a distribution of the differences in the averages for the four subjects common to the two experiments is made, the level of significance at which the mean is

different from zero is 13 percent for both the right and left ears. In other words, mean differences such as these could arise by chance 13 times in 100. It appears that the percent change in pitch at 200 cycles is greater than at 400 cycles; however it is possible that this difference could have arisen through chance variation.

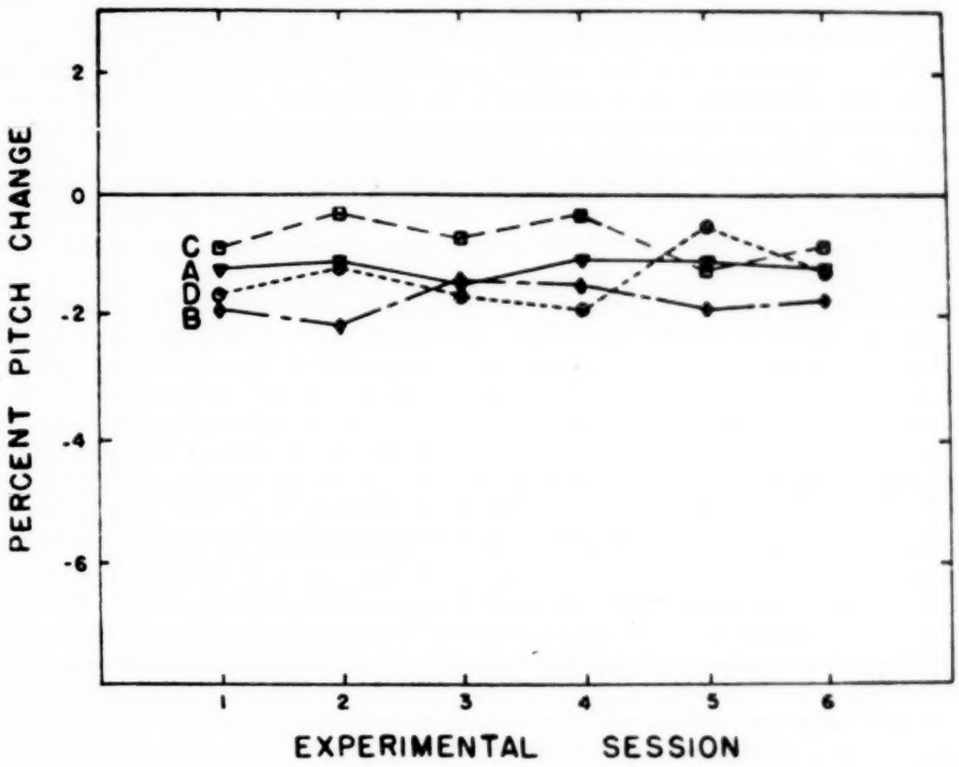


FIG. 3. Results at 400 cycles. The percent pitch change is shown for each subject at each of the experimental sessions.

EXPERIMENT 2

In the course of exploratory experimentation it was discovered that a loud tone in one ear would influence the pitch of a loud tone in the other ear even if these tones were not of the same frequency.[10] It thus appeared that the effect described in the first part of this paper might be only a special case of a more general effect of binaural interaction upon pitch perception. An investigation was therefore undertaken to determine the change in the pitch of a tone of given frequency when tones of various frequencies were introduced in the opposite ear.

APPARATUS

A separate circuit was used for each ear; each circuit consisted of an oscillator, attenuator, switch, and earphone. Rubber phone cushions were used on both earphones. The oscillator and monaural circuit used in Experiment 1 produced the tone of given frequency whose pitch was to be observed. For the production of the various other frequencies in the other ear a Western Electric 6010-B oscillator was used.

[10] These exploratory experiments indicated that the changes tended to disappear as the intensity was decreased below 95 decibels.

Procedure

Since many more frequencies were to be used in this experiment than in the previous one, a method of adjustment was used, in order to complete the determinations in a reasonable length of time. The affected tone, whose pitch was to be observed, was first sounded for .5 of a second; after a .5 second interval, this tone, together with an affecting tone in the other ear, was sounded for .5 of a second. Following a .5 second interval, this sequence was repeated once. If the subject found that the pitch of the affected tone changed when the affecting tone was sounded in the opposite ear, he signalled this fact to the experimenter. The affected tone was then sounded by itself every two seconds, and the subject adjusted its frequency (by means of the small frequency dial described in Experiment 1) so that its pitch sounded equal to the pitch which it had had when the affecting tone was simultaneously sounding in the opposite ear. By this method it was possible to determine the pitch shift in cycles in an affected tone produced by an affecting tone sounded in the other ear. If the subject noticed no pitch change, no adjustment was made. This method of adjustment was adopted in order to reduce the effects of fatigue to a minimum. The stimulus presentation-times are the same as those used in the first experiment. The repetition of the two tone sequence, described above, with equal intervals appearing between each tone, was adopted in order to eliminate time error.

The frequencies whose pitch was observed were 200, 400, 1000, 3000, and 4000 cycles. These frequencies, for clarity of exposition, will be designated as the *affected* frequencies. The frequencies introduced in the other ear, which will be designated as the *affecting* frequencies, were 120, 300, 500, 700, 1500, 2500, and 4000 cycles, and a frequency equal [11] to that of the affected tone. For determinations on the pitch of the 4000 cycle affected tone, an additional affecting tone of 4600 cycles was used. Affecting frequencies were chosen which lay far enough from the affected frequencies whose pitch was to be judged so that the subjects had no difficulty in judging the pitch of the affected tone to the exclusion of the pitch of the affecting tone. All tones used were at a loudness level of approximately 95 decibels. At this level the earphones were below the stage of overloading. The loudness levels were obtained in the same way as is described in Experiment 1.

A separate series of observations was made for each of the affected tones to determine the result of adding to the other ear the frequencies described above. Each such series consisted of two experimental sessions, separated by at least a day, at which five judgments were made of the pitch of the affected tone with each of the affecting tones in the other ear. At each experimental session the five judgments were divided into two groups, one of three, the other of two judgments. Adjustments with a given affecting frequency were divided in this way at a given experimental session, and between experimental sessions, in order to minimize the affect which one adjustment might have upon the other.

In all of these experiments the tone whose pitch was to be judged (the affected frequency) was introduced in the right ear, and the affecting tone was introduced in the left ear,[12] since preliminary experiments at 200 cycles did not indicate that the results differed appreciably when the tone whose pitch was to be judged was put in the left ear and the affecting tone was put in the right ear.

Results

For purposes of comparison, the results are plotted in Fig. 4 for each subject separately. The ordinate values of each graph represent the percent pitch change, in cycles, of a given affected tone when various affecting tones of the frequencies indicated on the abscissa are introduced in the opposite ear. Each point represents the mean of ten observations. The parameter values indicate the frequencies of the affected tones. Thus, the curve marked 200 shows the percent pitch change in a tone of 200 cycles caused by introducing a tone in

[11] Affecting tones equal in frequency to the affected tones were produced as in Experiment 1.

[12] However, for the case of an affecting tone equal in frequency to the given affected tone, determinations were also made with the affected tone in the left ear; the values obtained from the two procedures were averaged. This was done to prevent diplacusis from affecting the results.

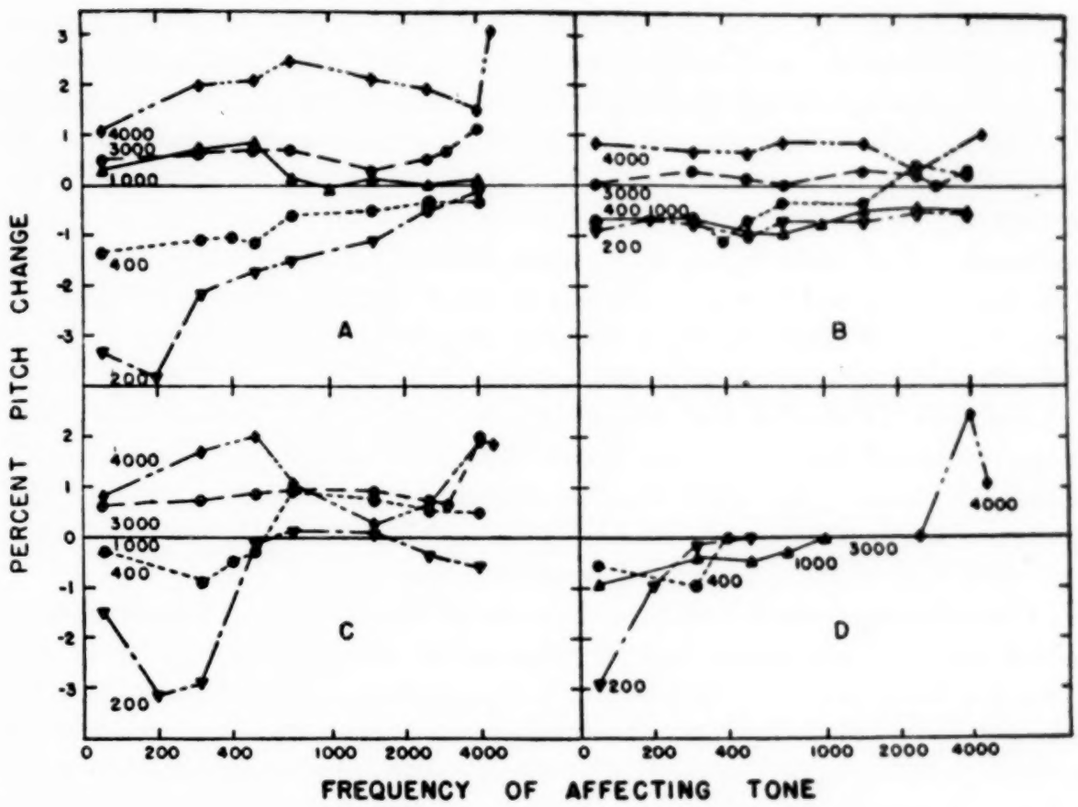


FIG. 4. The results of each of the four subjects are shown separately in graphs *A*, *B*, *C*, and *D*. The ordinate values of each graph represent the percent pitch change in cycles of a given affected tone when various affecting tones of the frequencies indicated on the abscissa are introduced in the opposite ear; the parameter values indicate the frequencies of the affected tones.

the opposite ear with frequency values indicated by the abscissa. Let us first consider subject *A*, whose results are shown in Fig. 4*A*. For the 200 cycle curve the pitch shift was in a downward direction; for the 400 cycle curve the pitch shift was still downward, but in a smaller amount. For 1000 cycles the pitch shift was slightly upwards; for the 3000 cycle curve the pitch shift was upwards to a greater degree; and for the 4000 cycle curve the pitch shift was upwards to the greatest extent.

The results of subject *B*, shown in Fig. 4*B*, are similar to those of Subject *A* in that the shift in pitch is downward for the lower affected frequencies, and upward for the higher affected frequencies;[13] however, the shift at 1000 cycles is downward, and there is no great difference between the percentage shifts for the three lower affected frequencies.

Subject *C* shows results for the affected tones similar to the first two subjects at the lower affecting frequencies, but shows some dif-

[13] For subject *B*, because of the presence of a marked diplacusis, no value was obtained at 4000 cycles for the pitch shift caused by introducing a tone of the same frequency in the opposite ear.

ferences from them at the middle and higher affecting frequencies. For example, for an affected tone of 200 cycles, the shift is downwards at the lowest affecting frequencies, but then becomes slightly upwards at the middle frequencies. For an affected tone of 400 cycles, this same effect is even more marked. At 1000 cycles, no pitch shift was detected. For the higher affected frequencies [14] the pitch shift is upward. But here again a peculiar effect, similar to that shown by the lower affected tones is shown, in that for the 4000 cycle affected tone the pitch shift upward for the middle affecting tones is below that for the 3000 cycle affected tone.

Subject *D* shows the most localized effects. Downward shifts were obtained for the three lowest affected tones only for the lower affecting tones. No shift was detectable for an affected tone of 3000 cycles. For an affected tone of 4000 cycles the shift was upwards; but only the highest affecting tones caused a shift.

Calculations were made to determine the levels of significance at which each of the mean values plotted in the graphs of Fig. 4 were different from zero. (Obviously, in those cases where no pitch change was reported, no calculations were made.) A *t* ratio (the ratio of the mean to the standard error of the mean) was computed for each mean value. Levels of significance were then determined from a table of the *t* distribution. It was found that 85 percent of the mean values were significantly different from zero beyond the 5 percent level; the remaining mean values (which were smaller numerically than the others) were different from zero at various higher levels of significance.

The results of all of the subjects are combined in Fig. 5. Curve *A* represents the average percent change in pitch caused by introducing a tone of the same frequency and intensity in the opposite ear. Curve *B* represents the average percent change in pitch caused by introducing tones of the same and different frequency in the opposite ear, all tones being at a 95 decibel loudness level; each value plotted is the average of the four values obtained from averaging each of the four subjects' results. The values in curve *B* at 200 and 4000 cycles are significantly different from zero at the 4 percent level; at 400 and 3000 cycles the values are significantly different from zero at the 19 and 7 percent levels respectively; furthermore, the values at 200 and 400 cycles are significantly different from each other at the 2 percent level, and the values at 3000 and 4000 cycles are significantly different from each other at the 11 percent level. The values of Curve *A* at

[14] With subject *C*, the values of the change in pitch for the higher frequencies were determined first for 4000 cycles, then for 3000 cycles, and then for 1000 cycles. This order was the reverse of that used with the other subjects. Subject *C*, however, showed upward shifts at 4000 cycles and 3000 cycles, and no shift at 1000 cycles. It can be concluded, therefore, that the order in which the determinations were made had no appreciable effect upon the direction of the pitch shift obtained.

200, 400, 3000, and 4000 cycles are different from zero, and from each other, at levels of significance similar to those for Curve *B*. From the foregoing values it can be concluded that the shape of the two curves of Fig. 5 is not determined by chance variation.

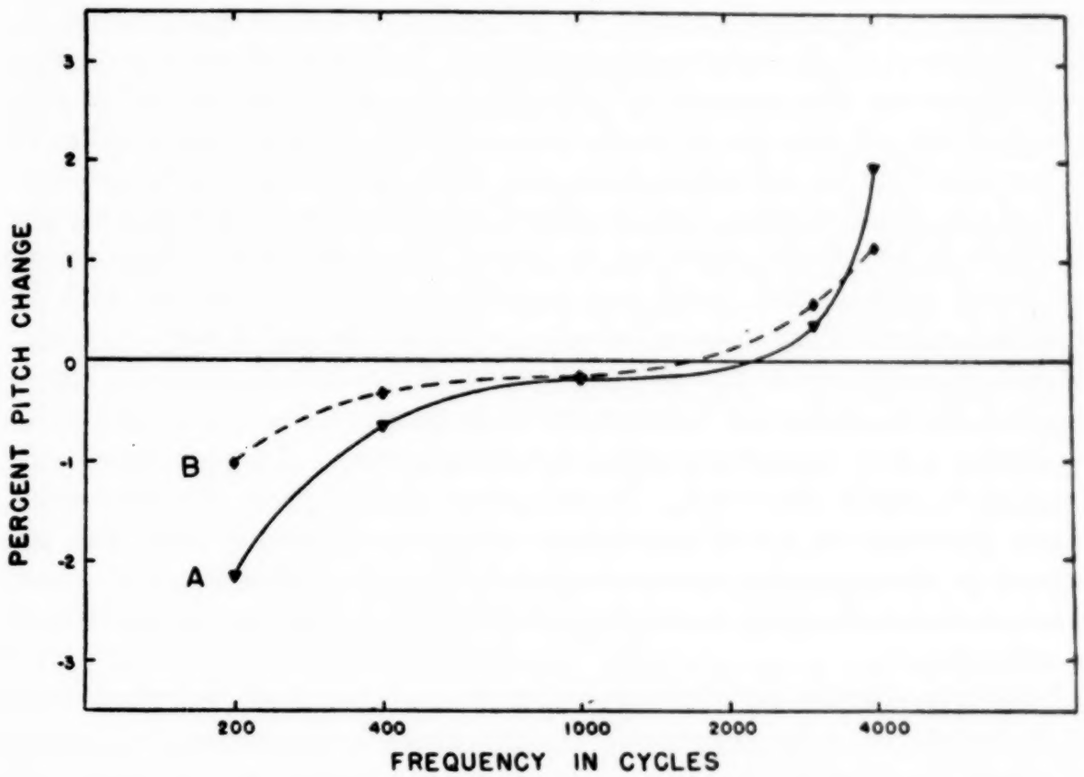


FIG. 5. Combined results of all of the subjects. Curve A represents the average percent change in pitch caused by introducing a tone of the same frequency and intensity in the opposite ear. Curve *B* represents the average percent change in pitch caused by introducing tones of the same and different frequency in the opposite ear, all tones being at a 95 decibel loudness level.

The effects of binaural interaction upon pitch found by the present experiment can be conveniently summarized from Curve *B*: Low tones are shifted, in general, downwards in pitch by an affecting tone of the same or different frequency introduced in the other ear; the shift of the 200 cycle tone is usually greater than that of the 400 cycle tone. (This same result was obtained in Experiment 1.[15]) High tones are shifted upwards in pitch by an affecting tone of the same or different frequency introduced in the other ear; the shift of the 4000 cycle tone is greater than that of the 3000 cycle tone. For tones of the middle range, 1000–3000 cycles, there is little if any change in pitch. The frequency at which the change in pitch is minimal varies within these limits according to the subject.

[15] The values obtained in Experiment 1 are slightly larger than those obtained in Experiment 2. This is probably due to the method used in Experiment 2, and also to the attitude of the subjects. (See Note 17.)

Control Experiments

1. *Binaural Peripheral Interaction*

It is of importance to know whether the change in pitch in one tone caused by adding another tone in the other ear is due to the influence of the component of the second tone which reaches the first ear by means of air or bone conduction. This possibility was checked by measuring the amount of threshold masking caused for a given affected tone in one ear by each of the affecting tones presented in the other ear.[16] The affecting tone was then introduced into the same ear as the affected tone, and at a level which caused the same number of decibels threshold masking as in the binaural case. Finally, the affecting tone at this level was combined in the same ear with the affected tone, which was now set at the 95 decibel level, and compared in pitch with the affected tone alone, which was at the 95 decibel level; the usual experimental procedure described under Experiment 2 was used for making the comparison. In no case was any change in pitch detected. It was thus shown that the changes in pitch observed in an affected tone when an affecting tone was presented in the opposite ear were not due to the influence of the component of the affecting tone which reached the other ear by air or bone conduction.[17]

An observation by Zurmühl (**27**),[18] together with his theoretical explanation of it, is of interest for the present discussion. Zurmühl observed that he was able to obtain greater shifts downward in the pitch of tones with an increase of intensity when he used binaural stimulation than when he used monaural stimulation. He explained this result by the hypothesis that when the monaural tone was sounded, the other ear was also excited, but to a much smaller extent; that this effect caused a higher pitch to be heard in this ear than in the first ear (Zurmühl dealt with low tones, which become higher in

[16] The amount of such masking was small; at some frequencies none could be detected. This is to be expected from Wegel and Lane's results on binaural masking (**24**). The use of rubber phone cushions in the present experiment undoubtedly made the amount of masking smaller.

[17] This procedure also provided a check on time error, since if such an error had been present, a difference in pitch might also have been expected. The absence of time error was due probably to the experience which the subjects had gained in making these judgments, and to the care with which the judgments were made. Experimental sessions were made short so that the subjects' attention did not wander during the session. It is of interest to note that two of the subjects, after the experiments were completed, reported that when they found that the pitch was changing in a given direction in an experimental session, they had often tried unsuccessfully to suggest to themselves that it was changing in the other direction; this, in order to prove to themselves that they were not making erroneous judgments. The absence of time error was also probably due to the method of presentation of stimuli (described in the Procedure of Experiment 2).

[18] Zurmühl's paper, with its reference to the early German work (especially that of Urbantschitsch), through accident was not obtained by the writer until after the present experiments were completed.

pitch with decreased intensity); and that the combination centrally of the two pitches caused a sensation intermediate between the two. It is unlikely, however, due to the small amount of stimulation reaching the other ear (as measured by masking in the present Experiment 2), that such stimulation in the other ear could affect the centrally perceived pitch to an extent sufficient to explain the subjectively marked pitch differences between monaural and binaural tones. In order to apply Zurmühl's hypothesis to the case where the pitch of a tone is affected by a tone of different frequency presented in the other ear, it would be necessary to assume that the only function of this tone introduced in the other ear is to mask the small component of the first tone present in that ear due to air or bone conduction. However, examination of the data of Experiment 2, as presented in Fig. 4, indicates that such a hypothesis will not explain them. The majority of the curves do not resemble masking curves. Some exploratory observations made on the change in pitch of the affected tone when the affecting tone is introduced in the same ear cast further doubt on this hypothesis. The observations were taken with affected tones of 200 and 4000 cycles; the regular series of affecting tones was used. Although the subjects found it difficult to make judgments because of the presence of beats and combination tones, they were able nevertheless to give consistent judgments. The change in pitch, in almost every instance, was in the same direction as in the binaural case, though somewhat smaller in extent. Irregularities indicated that probably other factors were entering in to complicate the results. Insofar as the results are similar to the binaural case, however, they indicate that a similar mechanism is acting even when the two tones are introduced in the same ear. Thus no simple explanation in terms of masking is possible. It can be concluded that Zurmühl's hypothesis is not adequate to explain the present results.

2. *Bilateral Muscle Action*

The possibility arises that a bilateral intra-aural muscle action may have produced the results. This is extremely unlikely, however, for it has been found that the action of the muscles in general decreases the intensity of the sound transmitted to the inner ear.[19] Yet in the present experiments, the changes in pitch obtained were in the same direction as has been reported by Stevens and Fletcher to result from an increase of intensity. That is to say, tones of low frequency shifted downward in pitch, tones of high frequency shifted upward in pitch, and tones of intermediate frequency shifted only to a negligible extent, if at all. Observations were made with the sub-

[19] Discussions of this literature are given in 17, 25, 26.

jects of the present experiment in order to determine the direction of the pitch shift with increasing intensity. The pitch of each of the affected frequencies was compared when sounded at a 40 decibel loudness level to that when sounded at the 95 decibel loudness level. In every case the change with increase in level was in the same direction as that reported by Stevens, and by Fletcher.

Observations were also made to determine the changes in pitch when yawning occurred, and when contractions of face and neck muscles simulating yawning took place, since it has been found that the intra-aural muscles contract under these conditions (**17**; **22**, p. 264). Changes in pitch at the lower frequencies did occur in some subjects, but were not consistent in direction even though all subjects reported a decrease in loudness under these conditions. No changes in loudness or pitch could be detected at the two highest frequencies used in the present experiments.

From the preceding considerations it does not appear that the results of the present experiment were caused by the action of the intra-aural muscles.

Discussion of Results

1. *Relation to a Principle of Maximal Stimulation*

It appears from the results of the preceding experiments that a change of pitch in a tone results from stimulation of the opposite ear with a tone of the same frequency; and also from stimulation of the opposite ear with tones of different frequency. A consideration of the experimental results, together with control experiments, indicates that the results are not due to a change in peripheral stimulation, caused by bone conduction or muscle action. It can therefore be concluded that the results are not produced by a change in the position of maximal stimulation on the basilar membrane.

Furthermore, it is difficult to see how a central interaction involving a shift in the position of maximal stimulation centrally could take place in a manner necessary to cause the results found in these experiments. For in the case where a tone of given frequency and intensity is altered in pitch by introducing a tone of the same frequency and intensity in the opposite ear, it is hard to see how an alteration of the position of maximal stimulation could occur such that the change in pitch is down for low frequencies, up for high frequencies, and nonexistent for frequencies of the middle range. The question of how this could occur becomes more difficult still when we consider the fact that a change in pitch in the same direction occurs when frequencies higher and lower than the given frequency are introduced in the other ear.

On the basis of the present results, then, it is reasonable to conclude that an adequate explanation of the mediation of pitch is not given by a principle of maximal stimulation.

2. *Relation to the Pitch-Intensity Function*

The direction of the pitch shift obtained in the present experiment is the same as that which Stevens and Fletcher obtained by increasing the intensity of the tones; that is to say, the pitch shifted downward for low frequencies, upwards for high frequencies, and to a negligible extent for a frequency region between. As has been mentioned in the discussion of control experiments, this agreement was confirmed by experimentation with the subjects of the present experiment. A further comparison with the results of Fletcher can be made, for it is a reasonable assumption that the effect of introducing a tone of given frequency and intensity in one ear, when a tone of the same frequency and intensity is sounding in the other ear, is to double the total loudness;[20] an examination of the curves published by Fletcher shows that the pitch change obtained in the present experiment is comparable to that which is produced through doubling the loudness at a 95 decibel level by an increase of the intensity. The values of the present experiment are slightly smaller than those given by Fletcher at the low frequencies, and slightly higher at the high frequencies.[21] The similarity of the results obtained in the present experiment to those obtained in investigations of the pitch-intensity function indicates that a similar mechanism is acting in both cases.

It is important to point out, however, that the results of Experiment 2 showed that the change in pitch of the affected tone was not due simply to an increase in the loudness of that tone itself. For it was shown in Experiment 2 that a change in the pitch of the affected tone could be produced by a tone of different frequency sounded in the other ear; that furthermore the magnitude of this pitch shift was comparable to that caused by an affecting tone of the same frequency introduced in the opposite ear. No systematic observations were made in the present experiment to determine what change in the loudness of the affected tone took place when the affecting tone was introduced in the other ear. A small number of exploratory experiments which the author made with himself as subject indicated that in many cases the loudness of the affected tone decreased when the affecting tone was introduced in the other ear; these results obtained

[20] Fletcher and Munson (13) obtained a loudness scale, constructed on the basis of this assumption, which agrees with scales obtained by other methods; thus the original assumption is validated.

[21] Since Stevens did not use levels of 95 decibels loudness, comparison with his results, which would involve an extrapolation of his curves, is not made.

when the affected and affecting tones were at the level used in the present experiment (95 decibel loudness level). If the affected tone was at a level near threshold, introducing the affecting tone in the other ear resulted in most cases in an increase in the loudness of the affected tone. This latter observation is the same as that reported by Urbantschitsch [22] in 1883 (**23**). He reported that a loud tone sounded in one ear would increase the loudness of a soft tone in the other ear; however, in a small number of tests with the same subjects, the opposite effect occurred. The variability of his results may have been partly due to the limitations of his tone sources, which consisted of watches and tuning forks. These results are undoubtedly related to those gained from experiments on binaural fatigue (**10, 18**), and binaural threshold summation (**14, 15**), as well as to the results of the present experiment—related in that they show a type of binaural interaction in which the effect of a stimulus in one ear is changed due to the presentation of a stimulus in the other ear.

The change of pitch of complex tones with an increase of intensity has been observed in some detail only recently (**11, 16**), although as early as 1883 Stumpf (**3**) reported observations on this phenomenon. The changes of complex tones evidently represent complicated examples of the affect of several frequencies upon the pitch of a given tone. However, another type of factor may enter into the effect, namely, the difficulty of abstracting the pitch of a fundamental from that of its overtones. A careful investigation is needed of the affects of one tone upon the pitch of another at high intensities for the case where the tones are closely related harmonically.

Conclusions from the Discussion of Results

Pitch, according to the results of the present experiment, is a function of binaural interaction. This binaural interaction produces effects similar to those produced by an increase in the physical intensity. Also it is in all probability related to other forms of binaural interaction in which the effect of a tone in one ear is altered by introducing a tone in the other ear. Furthermore, the dependence of pitch upon binaural interaction is of such a character as to cast grave doubt upon the hypothesis of the mediation of pitch by the position of maximal stimulation.

It seems possible that physiological experimentation, at higher intensity levels than have previously been used, will reveal more

[22] Urbantschitsch (see Note 18), also reported that the pitch of a tone in one ear was lowered when a loud tone of the same frequency was sounded in the other ear. That this latter observation did not strike him as being unusual is probably due to the fact that he was occupied at the time with various effects of interaction within a sense field rather than with the exact mechanism of pitch perception.

concerning the nature of this interaction; and continued psychophysical experimentation may be expected to define further the character of the function. The construction of an adequate theory must wait upon these investigations.

Summary

The pitch of a tone sounding in one ear can be changed by introducing a tone of the same frequency in the other ear, provided both tones are of fairly high intensity. The change in pitch, obtained in the present experiment with tones of a 95 decibel loudness level, is in the same direction as that which results from an increase in the physical intensity of the tone. Furthermore, the pitch of a tone can be changed by introducing a tone of different frequency in the other ear, provided, again, that both tones are of fairly high intensity. The direction of the change, obtained in the present experiment with tones of a 95 decibel loudness level, is almost without exception the same as that found to occur when a tone of the same frequency is presented in the other ear. The pitch of a tone is thus shown to be a function of binaural interaction; control experiments show that this interaction is not peripheral in nature.

The results of the present experiment do not support the hypothesis of the mediation of the perception of pitch by the position of maximal stimulation.

(Manuscript received July 23, 1942)

References *

1. Ringer, S. Über die Änderung der Tonhöhe bei Leitung von Tönen durch verschiedene Medien. *Pogg. Ann.*, 1863, **118**, 636–641.
2. Mach, E. Bemerkungen über den Raumsinn des Ohres. *Pogg. Ann.*, 1865, **126**, 331–333.
3. Stumpf, C. *Tonpsychologie*, Vol. 1. Leipzig: Herzel, 1883, 236–259.
4. Burton, C. V. *Phil. Mag.*, 1895, **39**, 447. (Original not available to the present writer; discussed in **9**.)
5. Broca, A. Influence de l'intensité sur la hauteur du son, *C. R. Sci.*, 1897, **124**, 1512–1514.
6. Broca, A. Influence de l'intensité sur la hauteur du son. *C. R. Biol.*, 1897, **10**, 652–654.
7. Bonnier, P. Pourquoi la tonalité d'un son perçu par l'oreille varie-t-elle avec son intensité? *C. R. Biol.*, 1897, **20**, 678–681.
8. Ewald, J. R. Eine neue Hörtheorie. *Pflügers Arch.*, 1899, **76**, 147–188.
9. Ewald, J. R., & Jäderholm, G. A. Die Herabsetzung der subjectiven Tonhöhe durch Steigerung der Objektiven Intensität. *Pflügers Arch.*, 1908, **124**, 29–36.
10. Bronstein, A. J. Sensibilization of the auditory organ by acoustic stimuli. *Bull. Biol. Méd. exp.*, 1936, **1**, 274–275, 276–277; **2**, 347–349. (Not available to the present writer; discussed in **22**, p. 220.)
11. Fletcher, H. Loudness, pitch, and the timbre of musical tones and their relation to the intensity, the frequency, and the overtone structure. *J. acoust. Soc. Amer.*, 1934, **6**, 59–69.
12. Fletcher, H. Newer concepts of the pitch, the loudness, and the timbre of musical tones. *J. Franklin Instit.*, 1935, **220**, 405–429.

* For convenience in citing, the references to the literature of the previous century on the pitch-intensity function are given together, number 1–9.

13. Fletcher, H., & Munson, W. A. Loudness, its definition, measurement, and calculation. *J. acoust. Soc. Amer.*, 1933, **5**, 82–108.
14. Holway, A. H., & Upton, M. On the psychophysics of hearing: III. The locus of the stimulus threshold. (Discussed in **22**, p. 52.)
15. Hughes, J. W. The monaural threshold: effect of a subliminal contralateral stimulus. *Proc. roy. Soc.*, London, 1938, **124** B, 406–420.
16. Lewis, D., & Cowan, M. The influence of intensity of the pitch of violin and 'cello tones. *J. acoust. Soc. Amer.*, 1936, **8**, 20–22.
17. Politzer A. *Diseases of the ear.* London: Bailliere, Tindall, and Cox, 1909, p. 66.
18. Rawdon-Smith, A. F. Experimental deafness. *Brit. J. Psychol.*, 1935, **26**, 233–244.
19. Rawdon-Smith, A. F. *Theories of sensation.* Cambridge Univ. Press, 1938. Section II: Audition.
20. Snow, W. B. Change of pitch with loudness at low frequencies. *J. acoust. Soc. Amer.*, 1936, **8**, 14–19.
21. Stevens, S. S. The relation of pitch to intensity. *J. acoust. Soc. Amer.*, 1935, **6**, 150–154.
22. Stevens, S. S., & Davis, H. *Hearing.* New York: John Wiley and Sons, 1938, Chapter 15.
23. Urbantschitsch, V. Über die Wechselwirkungen der innerhalb eines Sinnesgebietes gesetzten Erregungen. *Pflügers Arch.*, 1883, **31**, 280–309.
24. Wegel, R. L., & Lane, C. E. The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Phys. Rev.*, 1924, **23**, 266–285.
25. Wever, E. G., & Bray, C. W. The tensor tympani muscle and its relation to sound conduction. *Ann. Otol. etc.*, St. Louis, 1937, **46**, 947–961.
26. Wever, E. G. The electrical responses of the ear. *Psychol. Bull.*, 1939, **36**, 167.
27. Zurmühl, G. Abhängigkeit der Tonhöhenempfindung von der Lautstärke und ihre Beziehung zur Helmholtzschen Resonanz-theorie des Hörens. *Z. Sinnesphysiol.*, 1930, **61**, 40–86.

A QUANTITATIVE INVESTIGATION OF THE DELAY-OF-REINFORCEMENT GRADIENT [1]

BY C. THEODORE PERIN

Yale University

A number of animal learning experiments have been reported which involve a variable referred to as delayed reward or delayed reinforcement. This variable is usually characterized by a period of delay, spatial or temporal or both, between the occurrence of an instrumental response and the rewarding or reinforcing of that response. Examination of the results of the various experiments reveals striking differences in the degrees of learning which obtain with different lengths of reinforcement delay.

Watson (**12**) found that animals learn with equal facility to dig through a sawdust barrier to obtain food whether the food reward is given immediately upon successful completion of a trial or is delayed thirty seconds. Warden and Haas (**11**) found that a food-delay interval of five minutes does not retard maze learning. It is interesting to note that in both the Watson and the Warden and Haas experiments the animals were admitted to the feeding compartment immediately upon successful completion of a trial but that they were prevented from eating the food by the presence of a perforated lid covering the food dish. Experiments by Hamilton (**3**) and Wolfe (**15**) have demonstrated that animals learn a maze much less readily if they are detained *outside* the feeding compartment for one minute than if they are fed immediately on reaching the goal.

The fact that the Wolfe and Hamilton experiments obtained a decrement in learning with even a one-minute food delay is striking evidence of the presence of secondary reinforcement in the two earlier studies (**11**, **12**). In other words, in the Watson study and in that of Warden and Haas the feeding compartment and the presence of the food itself served as reinforcing agents of a secondary nature, to such an extent that the delay of the primary food reinforcement (eating) did not impair learning.

With the Wolfe and Hamilton studies there remains the possibility that the delay compartment, because of its spatial and temporal

[1] This investigation is a part of the coordinated research program of the Institute of Human Relations, Yale University. The writer is deeply indebted to Professor Clark L. Hull for generous advice and aid. The present paper is a portion of a dissertation presented to the Faculty of the Graduate School of Yale University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

proximity to the food compartment, will attain some strength as a secondary reinforcing agent, serving to reward an animal immediately upon completion of a run and thus reducing the effect of the delay period.

The question arises, then, as to how secondary reinforcement can be reduced to a minimum even if it is not possible to eliminate this factor entirely. With regard to the elimination of secondary reinforcement an experiment by Roberts (**9**) may prove to be significant. In this experiment the rats were required to touch a small wooden stick suspended in the problem cage. Following a delay period which varied in length for the different groups of animals, a door was opened giving the animal access to food in an adjacent compartment. Although there were too few animals in the different groups to yield a smooth gradient, Roberts found that under the conditions of his experiment a delay of 30 seconds in the reward materially reduces the rate of learning. A salient characteristic of the Roberts experiment consisted in having the response occur in the delay box itself. This removed the tendency for responses to be reinforced unequally in relation to their proximity to the delay chamber (secondary reinforcement).

A logical step from the Roberts study would be to have the response to be learned by an animal take place in the food box itself, using a single compartment as the problem box, delay compartment, and food box. With this procedure the animal would be allowed to make the response, the delay period would then begin, and at the end of the delay food would be given to the rat. In this way the stimulus environment of the animal would be much more constant during the delay interval, and the train of external stimuli leading to the food reward, characteristic of the maze situation, would be eliminated; this would materially simplify the problem of interpretation by eliminating an irrelevant factor.

Subjects and Apparatus

The subjects used in the present experiment were 116 male albino rats, approximately three months old. None of these animals had been used in any experiments previously. Three additional animals were discarded early in the procedure because they refused to eat in the experimental boxes.

The apparatus was similar to that used by Williams (**14**) and Perin (**7**), with certain modifications made necessary by the nature of the experiment. The experimental box consisted of a wooden outer shell about 21 in. high, 21 in. wide and 24 in. long. This shell was lined with a layer of Rock-wool insulation and inside this insulation was another wooden box measuring 12 in. along each of its three inside dimensions. A heavy wooden door with thick sponge-rubber lining was hinged to the outer shell, making the experimental compartment sound-proof to an extent that external noises apparently did not bother the animals. The box was ventilated by means of the laboratory suction line which served to remove the stale air and to draw fresh air into the compartment through several small openings. One modification of the apparatus was a smaller chamber on the floor of the 12-in. cubical compartment. This inner chamber was 6 in. wide, 5 in. high and 12 in. long. It was used in order to reduce the variability of the animals by confining them to a smaller space. A starting compartment in which the animal could be confined in order

to prevent its having access to the horizontal bar until the experimenter so desired was formed by a vertically sliding metal door between the entrance end of the interior compartment and the opposite end where the bar was located.

At the entrance end of the small interior chamber was a water bottle to which the animal had access at all times. At the other end of the compartment was a brass panel and through an aperture in this panel a small horizontal bar 55 mm in length could be inserted by the experimenter. Pressing this horizontal bar in order to obtain food was the task demanded of the animal. When the bar was retracted the aperture was covered by a sliding brass shutter. Spring tension on the bar was adjusted so that a downward pressure of 15 grams was needed in order to close a mercury-cup switch connected with the bar.

In series with the mercury switch was a magnetic signal marker mounted on a constant-speed, waxed-paper polygraph. Also in series with the mercury switch was a seven-volt relay which in turn operated a Telechron motor. Through a number of available circuits the latter could be used to operate a food-release mechanism at different time intervals after pressure on the horizontal bar had occurred.

The action of the apparatus through a single trial was essentially as follows: When the rat pressed the horizontal bar the response was registered on the waxed-paper record; the relay operating the 1 RPM Telechron motor simultaneously closed, and remained closed through the action of a holding relay. During one revolution a brass sweep on the Telechron motor made contact with five brass screws which were spaced so that the contact occurred 0, 2, 5, 10 or 30 sec. after the sweep started rotating. Each of these screws was wired to a selector switch which in turn was connected with the food-release mechanism. The selector switch could be set for 0, 2, 5, 10 or 30 sec. corresponding to the five screw contacts. Following the termination of the delay chosen by the experimenter, a pellet was released from the food magazine. After a pellet had been delivered, the motor completed its revolution and returned to the starting point where the circuit was broken automatically, setting the apparatus for the next trial.

A single food pellet served as one reinforcement. These pellets were cylindrical in shape, weighing about 0.05 gm. each. The composition of the pellets and the way in which they were made have been described elsewhere (7). The pellets were contained in a tubular magazine mounted, together with an attached food-release mechanism, on top of the box. When released, a pellet fell from the food magazine through another tube into the food tray directly under the horizontal bar in the experimental compartment.

Procedure

As mentioned previously, the learning of a bar-pressing habit with delay intervals of 0, 2, 5, 10, and 30 sec. between the occurrence of the response and the food reinforcement of that response was investigated. Five different groups of rats were run corresponding to the five different delay intervals used. Twenty-five animals were run in each of the first four groups and 16 animals were run under conditions of the 30-second delay.[2]

After being placed in individual cages the animals were fed Purina Dog Chow biscuits one hour each day for a week, in order to establish a definite feeding rhythm. On each of the two following days they were allowed to eat for an hour in the experimental boxes, the food being pellets of the same type as those used in the later reinforced trials.

Habituation to the Food-release Mechanism

Two days were then spent in training the animals to eat the pellets delivered to the food tray from the food-release mechanism. To do this an animal was first put into the experimental box with three pellets already in the food tray. While the animal was eating the third pellet the experimenter, watching closely through the glass window, closed a contact which caused the delivery device to drop another pellet into the food tray. As additional pellets were given the rats learned to respond to the click of the food mechanism by going to the food tray and eating the pellet which had been delivered. On the first day of this training each animal was given 30

[2] The desirability of having a 30-second delay group was seen only after the results had been obtained from the 100 animals in the four shorter delay groups. Hence, the 16 animals in this group were run quite apart from the other experimental groups, the conditions under which they were run duplicating the other groups with the exception of the longer food-delay interval used.

pellets and then was returned to its living cage for one-half hour; following this it was put back into the experimental box where it was given 30 more pellets. The day following, 60 more pellets were given to each rat in the same manner. Later discussion will point out reasons for the necessity of having this click-food-eating response very strongly fixated by the rats.

This habituation period also offered an opportunity to equate the experimental groups with regard to fear of the boxes and to learning ability. 'Freezing' (tendency to remain stationary in the box), defecation, urination, and failure to eat readily usually are indicative of a lack of adaptation to the experimental environment, and, in general, appear in combination. Another means of differentiating the animals is the rapidity with which they learn to obtain and to eat the pellets after the click of the food mechanism. Care was taken by the experimenter to have the animals which learned rapidly and the frightened animals, selected by these subjective judgments, mixed equally among the four experimental groups being run simultaneously in order to control possible group differences arising from a chance selection of these factors.

Training on the Bar-pressing Response

The day following 'habituation to the click' the animals were started on the bar-pressing training under the conditions of delayed food reinforcement. The rat was placed in the starting compartment of the experimental box, with the metal door blocking the approach to the horizontal bar. The heavy outer door was closed, confining the animal in the starting box. The polygraph was started, and when the experimenter saw that the rat was facing in the direction of the horizontal bar the door was raised by means of a string running into the interior compartment. This gave access to the horizontal bar and marked the beginning of the time interval elapsing before the animal pressed it. The raising of the confining door was recorded on the waxed paper by means of a magnetic signal marker operated by the experimenter.

In contrast to other studies involving the learning of a bar-pressing response, the present experiment did not make use of food smeared on or around the bar to attract the animal and to hasten the first responses, it being considered that the use of food in this way would distort the latency measure which was to be used as an indication of the strength of the bar-pressing response. The first, 'unlearned,' bar responses were made during the exploratory behavior of the animals in the vicinity of the bar and food tray. After an animal pushed the bar with sufficient pressure to operate the signal marker and to start the Telechron food-delay mechanism, the experimenter withdrew the bar by operating a lever outside the box.[3] Following the proper food-delay interval a pellet was dropped automatically into the food tray and was consumed by the animal.

It now becomes clear that the habit of getting the food from the tray immediately after delivery had to be very strong before this stage of training was started. There would be little point in having exact delay intervals between the response and the delivery of food if the interval between food delivery and food consumption varied considerably.

An animal was given 10 sec. after delivery of the food to eat the food. It was then returned to its living cage. The time between trials was eight min. for each rat. Fifty training trials were given to each of the animals in the five different delay groups.

Results

Equivalence of Groups

Equivalence of the experimental groups with regard to learning ability can be estimated by two methods: (1) the qualitative, subjective determinations by the experimenter during the habituation procedure (the basis upon which the group placement of the animals was made); and (2) the time required for the first bar-pressing response of the training series.[4]

[3] In the studies of Roberts (9) and Skinner (10), on the learning of an instrumental response under conditions of delayed reward, no provision was made for the prevention of recurring responses during the delay period.

[4] This was the only bar response by which the groups could be compared, since the different group delay periods began with the first response.

Although only one response from each of the 25 animals in a group makes a rather small sample for comparison, the closeness of the mean values and the very low critical ratios indicate that in all probability the groups are sufficiently well equated. The means of the time required for the first response are given, with critical ratios for the differences, in Table 1.

TABLE 1

EQUIVALENCE OF THE FIVE EXPERIMENTAL GROUPS ON THE BASIS OF THE LATENCIES FOR THE FIRST BAR RESPONSES OF THE TRAINING SERIES

Medians, means, standard deviations and critical ratios of the five groups indicate very little difference between groups for this measure.

	Group				
	0-Delay $N=25$	2-sec. $N=25$	5-sec. $N=25$	10-sec. $N=25$	30-sec. $N=16$
Median	23.5	24.0	22.5	25.0	27.5
Mean	31.6	29.8	35.8	35.8	36.6
σ	21.7	19.2	28.7	29.4	25.4
σ_M	4.3	3.9	5.7	5.9	6.4

Group	C.R.	P
0 vs. 2	.31	.38
0 vs. 5	.38	.36
0 vs. 10	.58	.29
0 vs. 30	.65	.26
2 vs. 5	.66	.26
2 vs. 10	.86	.20
2 vs. 30	.91	.19
5 vs. 10	.18	.42
5 vs. 30	.27	.40
10 vs. 30	.09	.46

The Effects of the Delay Periods

Practice curves for each of the four delay groups (Fig. 1) show clearly that the time between the occurrence of a response and the reinforcing state of affairs plays an important rôle in the learning process. These curves represent the median latency values for pooled successive sets of five trials of the animals in each experimental group, the latency being the elapsed time from the raising of the starting barrier until the rat pressed the horizontal bar. Table 2 gives the empirical values for these medians as well as the means of the distributions and the standard deviations.

The pooled practice curve for the immediate reinforcement (0-delay) group shows a very rapid drop in response time following a negatively accelerated decreasing growth function, and apparently becomes asymptotic to a value somewhat above zero latency (Fig. 1).

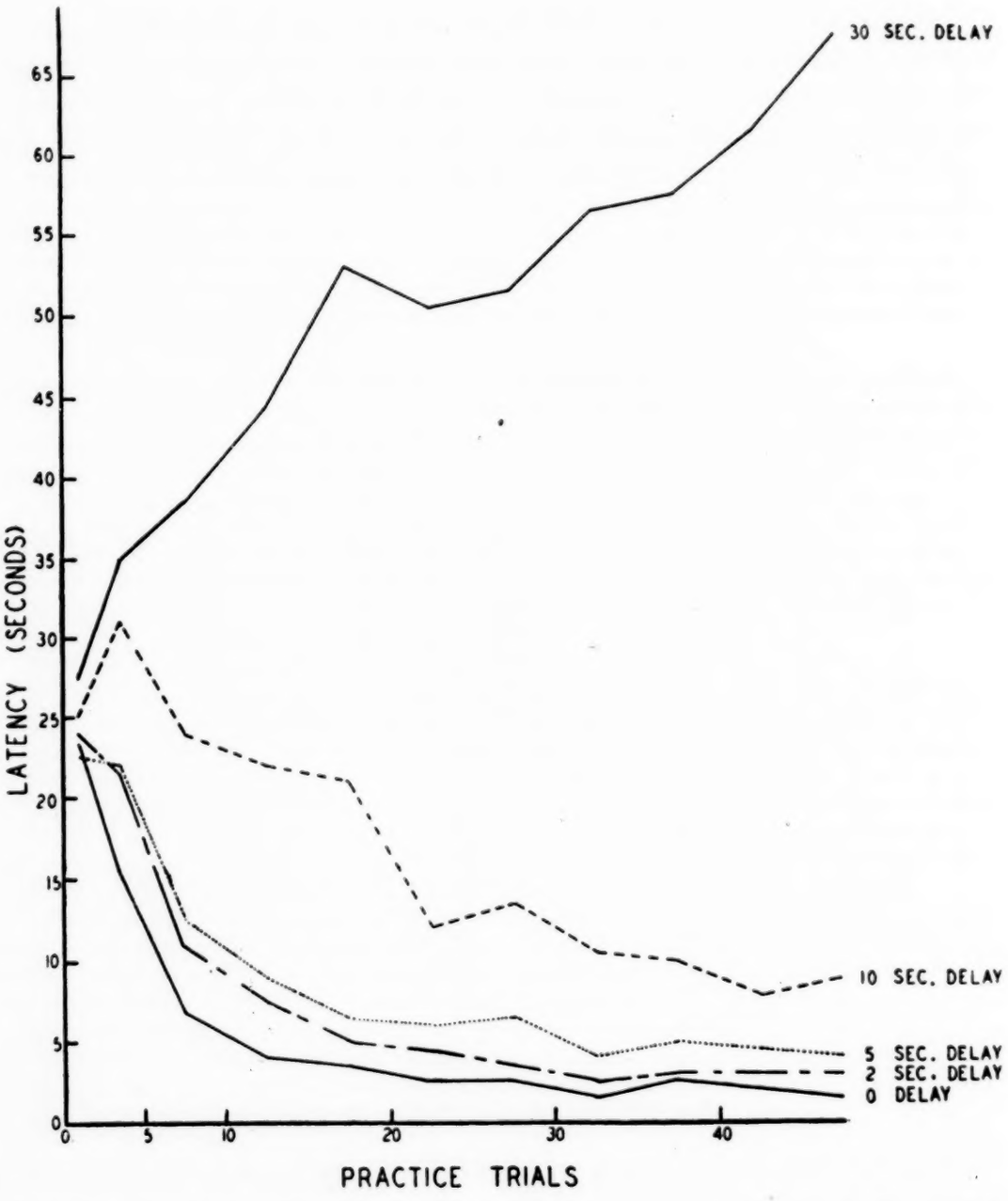


FIG. 1. Practice curves for the five experimental groups corresponding to the five conditions of delayed reinforcement. The curves for the 0, 2, 5, and 10-second delay groups are based on the results of 25 animals for each group, 16 animals comprising the 30-second group. Values represent the medians for pooled successive practice trials for all points except the first two points at the left of each curve. The first point on each curve represents the first trial only and the second point represents the median value for the 2–5 practice trials. Numerical values for the points are given in Table 2.

When a 2-second delay occurs between the response and the reinforcing state of affairs, the pooled curve for successive latencies resembles in form that for the 0-delay group. However, the rate of change appears to be somewhat less than that for the 0-delay group, and the median latency reached after 50 trials is longer than for the imme-

TABLE 2

MEDIAN AND MEAN RESPONSE-LATENCY VALUES (IN SECONDS) FOR POOLED DATA IN TERMS OF PRACTICE TRIALS FOR THE FIVE EXPERIMENTAL GROUPS

Standard deviations based on the means are also shown.

Practice Trials	Group														
	0-Delay			2-sec. Delay			5-sec. Delay			10-sec. Delay			30-sec. Delay		
	Md.	*M*	σ	Md.	*M*	σ	Md.	*M*	σ	Md.	*M*	σ	Md.	*M*	σ
1	23.5	31.62	21.69	24.0	29.78	19.70	22.5	34.34	28.72	25.0	35.84	29.39	27.5	36.61	25.4
2–5	15.5	21.40	15.59	21.5	22.38	12.24	23.0	34.68	21.64	31.0	45.28	28.84	35.0	49.50	34.0
6–10	7.0	11.13	7.07	11.0	14.10	9.07	12.5	23.63	26.70	24.0	32.36	24.10	38.5	60.51	39.2
11–15	4.0	10.24	12.29	7.5	10.53	6.21	9.0	16.08	18.10	22.0	29.40	28.30	44.5	64.00	38.0
16–20	3.5	8.24	10.31	5.0	8.53	4.41	6.5	10.60	10.11	21.0	28.78	16.90	53.0	66.00	44.0
21–25	2.5	4.97	6.25	4.5	6.37	4.32	6.0	10.23	7.75	12.0	16.62	11.60	50.0	64.79	40.5
26–30	2.5	4.36	6.64	3.5	5.19	3.36	7.0	9.69	6.53	13.5	17.73	9.70	51.5	68.00	38.0
31–35	1.5	2.62	2.84	2.5	3.84	3.03	4.0	5.71	3.25	10.5	14.78	7.30	56.5	70.01	44.7
36–40	2.5	3.64	2.72	3.0	3.53	1.74	5.0	7.02	3.59	10.0	14.07	9.40	57.5	74.00	48.2
41–45	2.0	2.46	1.71	3.0	3.82	2.11	4.5	6.07	3.90	8.0	10.49	6.60	61.0	72.26	51.4
46–50	1.5	2.14	1.54	3.0	4.20	2.10	4.0	5.46	2.77	9.0	12.73	8.10	67.5	79.63	54.2

diate-reinforcement group. In addition, the shape of the curve suggests that the latency under conditions of the 2-second delay may become asymptotic to a value substantially above the minimal latency attainable with the 0-delay group. Upon examining the practice curve for the 5-second delayed-reinforcement group it is evident that with this longer delay the slope of the curve becomes still more gentle, and that the latency after 50 reinforced trials is greater than that for the zero or the 2-second delay groups.

The 10-second delay group shows a still more gradual slope, and a higher minimum latency than the other three groups (Fig. 1). In addition, this group shows an interesting reversal or increase in latency early in the training, which is followed by the negatively accelerated decreasing growth function; this seems to be typical of the other three groups described previously. A possible reason for this reversal will be taken up in later discussion.

Unlike the four experimental groups having shorter delay periods, the 16 animals of the 30-second delay group failed to learn the bar response. Instead, the practice curve for this group (Fig. 1) demonstrates an increasing response latency as practice trials increase. Unfortunately, these animals were not given enough trials to determine whether this latency would continue to increase or whether it would approach some finite value as an asymptote. However, none of the animals showed any tendency for latency to decrease; *i.e.*, none demonstrated learning of the response.

Critical ratios were found for several cross sections of the practice curves for the 0–10 second delay groups (Fig. 1). These are pre-

TABLE 3

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN THE MEAN LATENCIES OF THE FOUR EXPERIMENTAL GROUPS HAVING 0–10 SECONDS FOOD DELAY, POOLED IN GROUPS OF SUCCESSIVE FIVE PRACTICE TRIALS, AT VARIOUS STAGES OF PRACTICE

Delay Groups	Practice Trials			
	21–25	31–35	41–45	46–50
0–2	.92	1.47	2.52	3.96
0–5	2.63	3.59	4.20	4.68
0–10	3.94	7.79	5.90	6.42
2–5	2.18	2.10	1.62	1.83
2–10	4.15	6.92	4.80	4.11
5–10	2.29	5.67	2.89	4.25

sented in Table 3. The significances of the differences between the groups at the several points may be taken as rough indication of the discreteness of the four curves in general. However, the fact that the curves are fairly smooth and that there is practically no overlapping is perhaps the best indication that there is a real rather than chance difference between the curves.

In summarizing the empirical results, the following statements can be made: The longer the time interval between the occurrence of the bar-pressing response in the present situation and the reinforcing of that response by food reward, the slower will be the rate at which the response is learned as measured by the latency of the response. Also, the longer the food-delay interval, the less will be the maximum habit strength which can be attained under the given conditions.

DISCUSSION

In order to see in detail what effects the delay-of-reinforcement interval has on the learning process, mathematical curves were fitted to the empirical practice curves shown in Fig. 1. With the immediate reinforcement group, the practice curve follows the equation:[5]

$$L = 18.58 \cdot 10^{-.0614T} + 1.85, \qquad (1)$$

where L is the latency or response time and T is the number of training trials. In this equation the value 1.85 is a constant representing the presumptive asymptote of the 0-delay group, 18.58 is a constant, and $-.0614$ is an exponential constant determining the rate of decrease for latency as the number of training trials increases.

The equation describing the curve for the 2-second delay group is:

$$L = 27.44 \cdot 10^{-.0568T} + 2.75, \qquad (2)$$

[5] The following problems in curve fitting were solved by Mr. Bengt E. Carlson.

The equation for the 5-second group is:

$$L = 23.30 \cdot 10^{-.0534T} + 4.10. \quad (3)$$

The smoothness of the curve for the 10-second delay group is disturbed somewhat by the initial increase in latency which is followed by a negatively accelerated decreasing function. The curve fitted to the date for this group does not include the first two points but instead begins with pooled trials 6–10. The equation of this curve for the 10-second delay group is:

$$L = 32.52 \cdot 10^{-.0312T} + 7.45. \quad (4)$$

It was pointed out in the statement of the results that the ultimate limit of habit strength as indicated by response-latency is a function of the size of the reinforcement delay interval. Examination of equations 1–4 shows this functional relationship in terms of the limits of the curves as determined by extrapolation. The values of these habit strength limits are given in Table 4. It must be remembered

TABLE 4

LATENCY ASYMPTOTES, RECIPROCALS OF THE ASYMPTOTES AND EXPONENTIAL CONSTANTS FOR THE FOUR EXPERIMENTAL GROUPS HAVING 0–10 SECONDS' FOOD DELAY

Values were obtained from equations fitted to the practice curves for the four groups.

	0-Delay	2-sec.	5-sec.	10-sec.
Latency asymptote (extrapolations based on practice curves)	1.85	2.75	4.10	7.45
Reciprocal of latency asymptote	.541	.364	.244	.134
Value of exponential constant	.0614	.0568	.0534	.0312

that latency is related in an inverse manner to most other measures of habit strength; that is, the stronger the response tendency, the shorter the latency. In order to transform the latency results into values more easily compared with other measures of habit strength, the reciprocals of response latency, constituting a positive function of habit strength, have been determined. These reciprocals are given in Table 4, and Fig. 2 is a plot of these values.

Two fundamental questions with regard to both the gradient of reinforcement and the delay-of-reinforcement gradient have been concerned with the actual shape of the gradients (whether logarithmic, straight line, exponential, etc.) and with the probable active limits. A curve was fitted to the reciprocals of the response-latency asymptotes in order to obtain a mathematical equation describing the gradient obtained in the present experiment. A very good fit to the

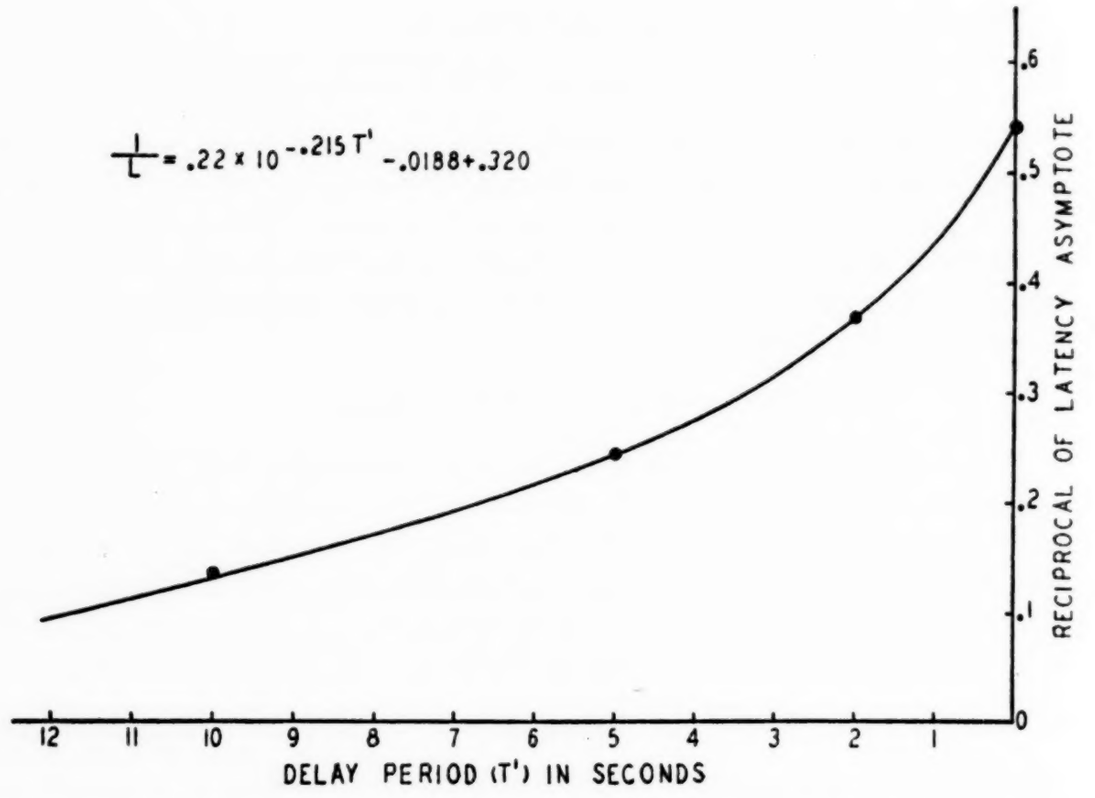


FIG. 2. Plot for the reciprocals of the four latency asymptotes for the four experimental groups having 0–10 seconds' food delay. The reciprocal values are shown as solid dots and the connecting curved line represents a theoretical function fitted to the points, the function corresponding to the equation $1/L = .22 \cdot 10^{-.215T'} - .0188T' + .320$. Numbers shown on the abscissa are reversed to represent time along a continuum running from left to right with the responses occurring at different intervals of time before reinforcement (R) at right of abscissa. Values for the points are given in Table 4.

empirical values was obtained, the theoretical curve being presented with the empirical points in Fig. 2. The equation for this curve is:

Equation (5) $1/L = .22 \cdot 10^{-.215T'} - .0188T' + .320$.

Mathematically this fitted curve is an additive function, and it is best described as an exponential function approaching a sloping straight line as an asymptote.

If we look again at equations 1–4 another trend becomes obvious. This is in the exponential constants determining the rate of decrease for latency. The values of the constants showing this trend in the four experimental groups are presented in Table 4 and Fig. 3.[6]

Examination of Fig. 3 discloses that the exponential constants have a decreasing trend as the interval of delay between the responses and reinforcement increases. These constants are subject to a certain degree of variability because of the variability of the original data and

[6] The results of the 30-second delay group are not included in this quantitative analysis because of difficulties in determining the ultimate limit of the practice curve for this group from the data available.

the chance errors arising from curve fitting; it would therefore be unwise to say that the values represent the final function which would be obtained from an unlimited sample. However, there is a definite downward trend indicated by the exponential values. Rather than attempt to fit a single curve to these points and then extrapolate to the X-intercept in order to determine at what delay period the exponent would have a zero value, two functions which were considered to represent extreme possibilities were fitted to the points. The first function involves the assumption that the trend of the points can be described best by a straight line (in spite of a definite suggestion of curvature in the trend). This straight line function, Fig. 3, corresponds to the equation: $Y = .0634 - .00298T'$, where Y is the exponential constant and T' is the delay interval in seconds.

On the other hand, the trend of the points in Fig. 3 suggests that a simple curvilinear function might be more suitable than a straight line, and for this reason a simple parabola was also fitted to the points. The equation for this function is: $Y = .06 - .00029\ T'^2$, where Y is the exponential constant and T' is the delay interval.

When the straight line function and the simple parabola in Fig. 3 are extrapolated in order to find where they cross the base line it is

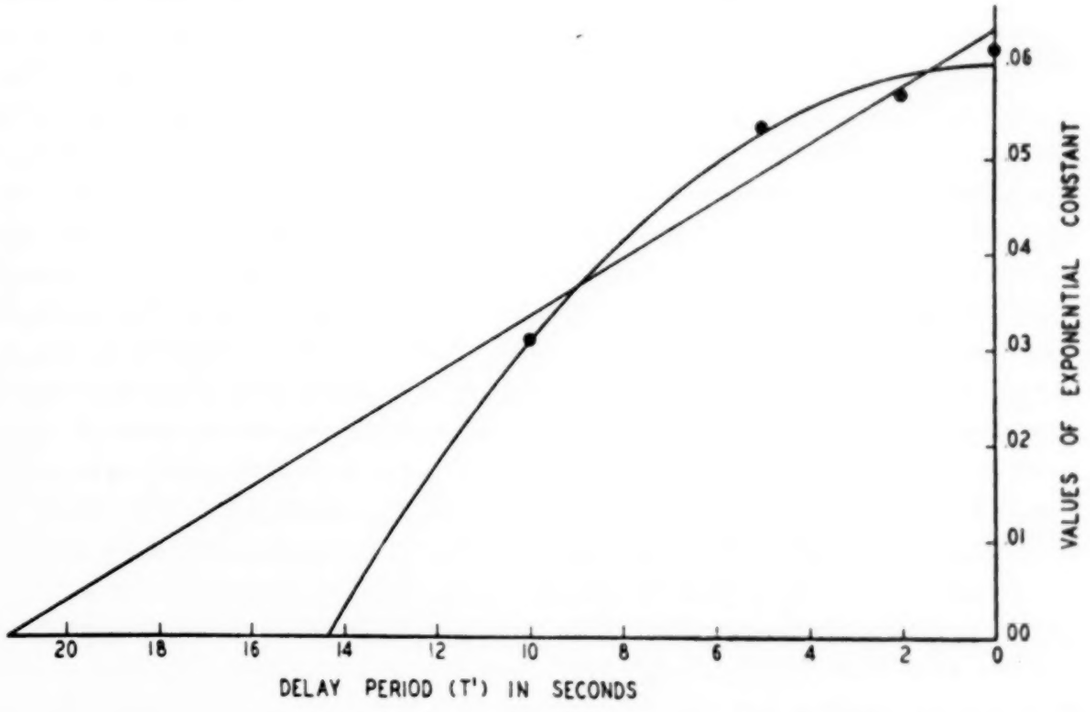


FIG. 3. Showing the values (solid circles) of the exponential constants in the equations of the practice curves for the four experimental groups having 0–10 seconds' food delay. The straight line fitted to the four points corresponds to the equation $Y = .0634 - .00298T'$ and the curved line represents a simple parabola fitted to the points and corresponds to the function $Y = .06 - .00029T'^2$.

found that the X-intercept is at 21.3 sec for the straight line and at 14.4 sec for the parabola. Subject of course to experimental verification, these values suggest that between 14 and 21 sec is the value beyond which there can be no further delay if the bar response is to be learned by primary reinforcement.

The curve based on reciprocals of the latency asymptotes shown in Fig. 3 is consistent with this conclusion. Equation 5 indicates that this curve may reach a value of zero when the food delay interval is about 17 seconds, the zero value indicating that with this delay an infinite latency may be reached with continued training on the bar response.

The results of the 30-second delay animals indicate that this amount of food delay is beyond the limit which will permit the response to be learned. Also, five additional animals were run under conditions of one-minute food delay. Of the five rats one 'extinguished' after making six responses of gradually increasing latency. The other four animals all had ceased to make bar-pressing responses by the 15th trial. Generally extinction was characterized by the animal's going to sleep in the experimental chamber. There is a question as to the proper use of the word 'extinction' in this instance, since it is customarily applied to the decrement in habit strength found with successive unreinforced evocations of a learned response. However, an element of extinction can be identified in the situation. Before the rats were given trials on the bar-pressing response, they were trained to eat pellets which were released into the food tray. Later, when the bar was present above the food tray, this previous training served to make the animal approach the bar and directed the animal's activity in the vicinity of the bar. In all probability this increased the chance for a bar-pressing response to occur through random behavior in the initial stages of learning. When the animals were being trained to press the bar, however, the experimenter no longer released pellets into the food tray and therefore the response of crossing the box to the tray would be subject to extinction if there were no other sources of reinforcement. Food pellets delivered after the delay period following the bar-pressing response might then be considered to have reinforced the response of crossing the box toward the food tray, in addition to reinforcing the bar response itself.

The writer believes that this mutual dependence of the bar response and the response to the food tray accounts for the initial increase in latency for the 10-second delay group as shown in Fig. 1. Presumably this increase in latency is halted by the fact that even with a 10-second delay the increment of reinforcement for each trial is sufficiently large for the response to be learned and for the latency of the bar response to decrease eventually.

The delay-of-reinforcement gradient found in the present experiment (Fig. 2) appears to be contradictory, in at least one respect, to the gradient of reinforcement hypothesis advanced by Hull (**6**). The present delay gradient not only approaches a zero base line as the interval of delay is increased, but indications are that it actually intersects this line at a delay considerably less than 30 seconds. This intersection has been interpreted as the point beyond which rats, on the average, are unable to learn the bar-pressing response. On the other hand, Hull (**6**) has postulated a gradient of reinforcement which approaches the zero base line but becomes asymptotic to the line, which indicates that, however remote the response from the reinforcement, there would be a positive increment of reinforcement for the occurrence of that response. However, the writer is of the opinion that the present results do not negate in any way the gradient-of-reinforcement in the form postulated by Hull. The delay-of-reinforcement gradient as presented in Fig. 2 is in terms of the time required for the response to occur, response latency in this case being considered an indirect measure of effective habit strength. But in addition to the factor of the bar-pressing habit resulting from previous reinforcement, probably there are also factors of accumulated inhibition resulting from bar-pressing responses as well as the reinforcement or extinction of the animal's response of crossing the box to the food tray.[7] In addition, there are probably other positive and negative tendencies which have not been identified here. Presumably the final act of bar pressing is the resultant of the complex interaction of all these tendencies. As the present writer understands the gradient-of-reinforcement hypothesis in its application to this situation, it represents the accumulated increments of habit strength of the bar response resulting from previous reinforcement. It may well be that the pure gradient is an exponential function approaching zero habit strength as an asymptote, *i.e.*, increments of reinforcement are obtained even with extremely long delays; but that the additional factors of extinction and inhibition in this bar-pressing situation with long delay periods become strong enough to override the response tendency and thus no response occurs.

Since indications are that learning could not occur if primary reinforcements were delayed as long as 20–30 seconds in the absence of secondary reinforcement factors, it would seem that a premium is placed on secondary reinforcement in the learning of responses remote from the reinforcing state of affairs. If different aspects of the stimulus situation, other than the primary reward stimulus, and the responses to these stimuli assume the power of reinforcement, then it is

[7] This probably involves the same factors as the 'inhibition of reinforcement' phenomenon reported by Hovland (**4**).

possible for responses occurring quite outside the range of primary reinforcement to be learned. Hull (**6**) has presented a theoretical account of the action of secondary reinforcement in mediating the learning of responses too remote from the primary reinforcement to be learned without the presence of the secondary reinforcement stimulus characteristics.

Summary

1. Four groups of albino rats, 25 animals in each group, were trained on a bar-pressing response. With one group the food pellet reward was given simultaneously with the occurrence of the response. A second group received food after a delay of 2 seconds following the bar response. The third group was given food after a 5-second delay, and the fourth group after a 10-second delay. A fifth group made up of 16 animals received training with the food reinforcement being delayed 30 seconds following the bar response.

2. Fifty reinforced trials were given to the animals in each group. The measure of habit strength was response latency (the time required for the response to occur after an animal had been given access to the bar).

3. Since latency is an inverse function of habit strength the practice curves show a decreasing trend as learning progresses. The rates at which the practice curves decrease show a consistent relation to the length of the delay period. Animals reinforced simultaneously with the response demonstrate the most rapid learning, with the rate of learning becoming increasingly slow as the delay interval is increased.

4. The practice curves are described by means of equations fitted to the empirical values. Indications are that the minimum latency which ultimately will be reached with extensive practice is a function of the length of the delay period preceding reinforcement: the longer the delay, the less will be the habit strength ultimately reached with a maximum number of reinforced trials.

5. Reciprocals of the response-latency asymptotes for the practice-curve equations appear as a negatively accelerated decreasing function of the length of the delay-of-reinforcement interval. The equation fitted to the reciprocal values indicates that learning does not occur if the reinforcement is delayed longer than 17 seconds. It is recognized, however, that the range of extrapolation is too great to fix this limiting time interval accurately.

6. Two equations are given which describe the exponential constants for the practice curves as functions of the delay interval. Indications are that within the range of 14–21 seconds' delay the exponential constant probably reaches zero. This is considered to

indicate that on the average learning cannot take place with a delay interval larger than 14–21 seconds, and is consistent with the findings for the asymptote reciprocals.

7. A possible explanation is given for the differences between the delay-of-reinforcement gradient obtained in the present experiment and the gradient of reinforcement postulated by Hull.

8. It is believed that the present experiment has been more successful in eliminating factors capable of acquiring the power of secondary reinforcement than have earlier studies, and that this probably accounts for the much shorter delay-of-reinforcement gradient than previous studies have indicated. The fact that elimination of factors of secondary reinforcement results in an extremely short gradient is taken as evidence that in most learning situations secondary reinforcement plays a prominent role.

(Manuscript received July 1, 1942)

References

1. Bugelski, R. Extinction with and without subgoal reinforcement. *J. comp. Psychol.*, 1938, **26**, 121–134.
2. Grice, G. R. An experimental study of the gradient of reinforcement in maze learning. *J. exp. Psychol.*, 1942, **30**, 475–489.
3. Hamilton, E. L. The effect of delayed incentive on the hunger drive in the white rat. *Genet. Psychol. Monogr.*, 1929, **5**, 131–207.
4. Hovland, C. I. 'Inhibition of reinforcement' and phenomena of experimental extinction. *Proc. nat. Acad. Sci.*, 1936, **22**, 430–433.
5. Hull, C. L. The goal gradient hypothesis and maze learning. *Psychol. Rev.*, 1932, **39**, 25–43.
6. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century Co. (in preparation).
7. Perin, C. T. Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93–113.
8. Perin, C. T. A quantitative investigation of the delay-of-reinforcement gradient. Ph.D. dissertation, Yale University, 1942.
9. Roberts, W. H. The effect of delayed feeding on white rats in a problem cage. *J. genet. Psychol.*, 1930, **37**, 35–58.
10. Skinner, B. F. *The behavior of organisms.* New York: D. Appleton-Century Co., 1938.
11. Warden, C. J., & Haas, E. L. The effect of short intervals of delay in feeding upon speed of maze learning. *J. comp. Psychol.*, 1927, **7**, 107–116.
12. Watson, J. B. The effect of delayed feeding upon learning. *Psychobiol.*, 1917, **1**, 51–60.
13. Williams, K. A. The reward value of a conditioned stimulus. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 31–55.
14. Williams, S. B. Resistance to extinction as a function of the number of reinforcements *J. exp. Psychol.*, 1938, **23**, 506–522.
15. Wolfe, J. B. The effect of delayed reward upon learning in the white rat. *J. comp. Psychol.* 1934, **17**, 1–21.

RETENTION OF NONSENSE SYLLABLES IN INTENTIONAL AND INCIDENTAL LEARNING

BY WILLIAM C. BIEL AND RONALD C. FORCE

Miami University

Results from relevant previous experiments [1] agree in showing that material which is learned with the intent to learn or with the intent to remember is retained better than material which is learned without such intent to learn to remember (learned incidentally). In most of these experiments, however, there has been no attempt to equate the amount of original learning which has taken place in the two groups; instead, it has been the number of presentations of the material to be learned that has been held constant. The latter method, therefore, does not control the possibility that the group with intent to learn probably learns more originally than does the other group, and as a result of this retains more at a later time.

In an experiment such as Peterson's (3), however, an attempt has been made to take this factor into account. He compared two groups in the percentages of the material immediately recalled with the percentages obtained in later recall. He found that on immediate recall the group with intent to learn was from 15 to 30 percent better than the group without intent and was approximately 50 percent better on delayed (48 hours) recall. Although this method points out general trends, comparisons nevertheless are more accurately made when the original amount of learning is equal.

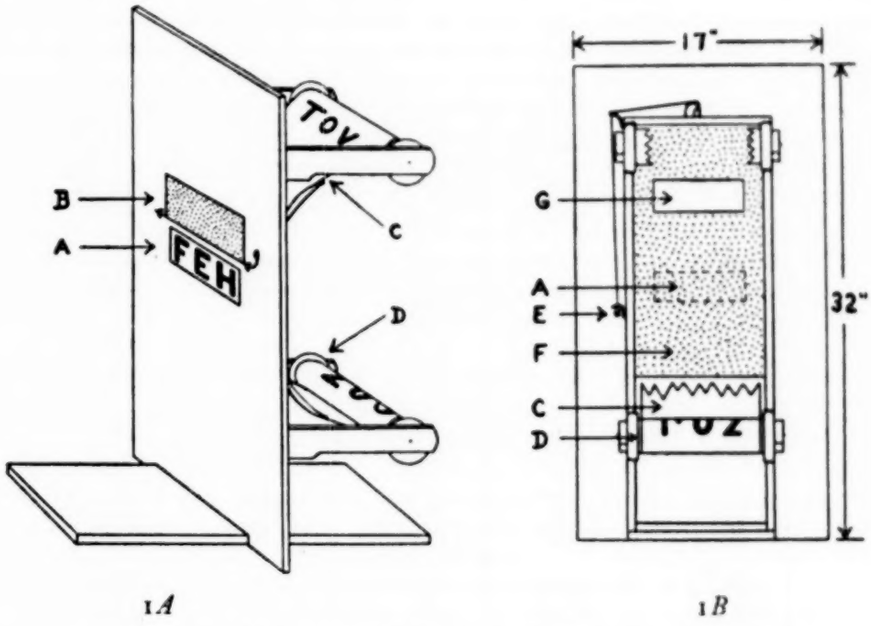
The present experiment is an attempt to determine the effect of intentional and unintentional (incidental) learning upon the retention of nonsense syllables, *when the amount of original learning is equal.* In order to make the amount of original learning equal in the two groups, the individuals in the non-intent group were given the necessary number of trials to make the average number of nonsense syllables immediately recalled the same as that for the intent group. This procedure will be described in detail in the next section.

Apparatus and Procedure

1. *Tachistoscope.*—A drop tachistoscope was used for presenting a series of nonsense syllables to the subjects in a classroom situation. A diagram containing the essential dimensions of the tachistoscope is shown in Fig. 1. An exposure time of one-thirtieth of a second for each syllable was used because this interval had been found in preliminary experimentation to be long enough

[1] For review of the literature see Thisted and Remmers (4).

for adequate perception of the syllables at the various distances and angles from which they had to be viewed by the subjects. The time was short enough, however, to promote in one group the intended attitude (to be described later) that the purpose of the experiment was to test legibility of printing types.



1*A* 1*B*

FIG. 1. Tachistoscope used for presenting nonsense syllables.

Twelve nonsense syllables were exposed individually through a 6″ × 1⅛″ aperture labeled *A* in Fig. 1*A*. The letters were 1½″ in height and each syllable had been drawn [2] on a separate sheet of white paper which in turn had been glued to an oilcloth strip (*C*). This strip was fastened to, and stretched between, two rolls, thus providing a continuous mounting. Exposures were brought about by releasing the tab, *E* (shown in Fig. 1*B*), which in turn released the drop-shutter, *F*. This drop-shutter, which was black and had a two-inch opening, *G*, dropped past the aperture *A* of the front of the tachistoscope and exposed momentarily the nonsense syllable which was to the rear of both openings. Any one part of the syllable was exposed for one thirtieth of a second as the shutter dropped by. The next syllable was then rolled into position by turning the bottom roll, *D*. The front shade, *B*, was then lowered to prevent exposure of the next syllable as the drop-shutter was returned to the 'ready' position at the top of the slide as shown in Fig. 1*B*. The time interval between successive presentations was fairly constant and was approximately 13 seconds in length.

2. *Nonsense Syllables.*—Syllables from Glaze's list (1) having zero percent association-value were used as learning material. A list with 12 syllables was used because it was possible with that number to get the average recall scores of the groups up to 6 or 7 in a practical experimental period. The syllables were so selected that no two first letters of any of the syllables were the same, and 6 vowels were used twice so that no one letter was predominantly used in the syllables. The syllables used are shown in Table 1.

3. *Order of Syllables in the Lists.*—The order of the syllables in the list of 12 nonsense syllables was mixed in successive presentations of the list to the groups. This was necessary in order to foster the attitude in the non-intent group that this was an experiment on legibility; the same mixed order was followed in the presentations to the intent group.

The order in which the syllables were presented is shown in Table 1. In planning the lists, the experimenters kept in mind the fact that position in a list of syllables will influence the retention value of a syllable. The syllables were so placed that no one syllable had more positions favorable for learning (early or late in the list) than any other syllable.

[2] The letters were printed by projecting small point types of the faces used, tracing the image projected, and then filling in the outlines solidly with India ink.

TABLE 1

ORDER OF SYLLABLES IN LISTS, AND NAMES OF PRINTING TYPES IN WHICH THEY APPEARED IN THE TACHISTOSCOPE

Position Number	List Number					
	1	2	3	4	5	6
	Syllable (Type)	Syllable (Type)	Syllable (Type)	Syllable (Type)	Syllable (Type)	Syllable (Type)
1	TOV (K)	VYF (K)	GIW (SR)	JYC (B)	WUH (OS)	ZOJ (SR)
2	FEH (COS)	GIW (OS)	VYF (B)	WUH (COS)	JYC (SR)	LAJ (OS)
3	WUH (SR)	ZOJ (B)	LAJ (Ch)	QUG (COS)	CEF (Ch)	DAX (K)
4	JYC (Ch)	LAJ (COS)	ZOJ (COS)	CEF (B)	QUG (K)	SIJ (B)
5	GIW (K)	JYC (OS)	WUH (Ch)	ZOJ (Ch)	LAJ (SR)	QUG (OS)
6	SIJ (Ch)	TOV (COS)	FEH (K)	GIW (Ch)	VYF (OS)	WUH (B)
7	QUG (B)	SIJ (SR)	DAX (B)	FEH (OS)	TOV (B)	VYF (Ch)
8	VYF (COS)	WUH (K)	JYC (K)	LAJ (K)	ZOJ (K)	CEF (K)
9	ZOJ (OS)	CEF (SR)	QUG (SR)	SIJ (K)	DAX (COS)	FEH (SR)
10	CEF (OS)	DAX (Ch)	SIJ (OS)	TOV (SR)	FEH (Ch)	GIW (COS)
11	LAJ (B)	QUG (Ch)	CEF (COS)	DAX (OS)	SIJ (COS)	TOV (Ch)
12	DAX (SR)	FEH (B)	TOV (OS)	VYF (SR)	GIW (B)	JYC (COS)

Code for printing types: B = Bodoni; COS = Caslon Old Style; Ch = Cheltenham; K = Kennerly; OS = Old Style; SR = Scotch Roman.

The lists were presented in the order 1, 2, 3, 4, 5, 6, 5, 6, 4, 3, 2, 1 to the non-intent group and in the order 1 through 5 to the intent group. One group of 29 subjects in the intent group, however, was shown lists 1 through 6 in that order. The number of times that it was necessary to present the lists to the non-intent and the intent groups in order that the original learning of the syllables be equal had been determined by preliminary experimentation. The findings from the preliminary experimentation indicated that 12 presentations of the lists to the subjects in the group without the intent to learn produced the same average immediate recall as did between 5 and 6 presentations of the list to the intent-to-learn group. In order to achieve comparable averages in the intent and non-intent groups in the main experiment, therefore, a part of the intent group was given 5 presentations of the list and the remainder of that group, 6 presentations. Although individual subjects in the two groups were to be matched for the number of syllables recalled, it was important that the two total groups averaged approximately the same so that a

selection of subjects in one group better able to memorize was not being matched with subjects in the other group who were less able to memorize.

4. *Printing Types.*—Six printing types were used. Five of them were selected from those which Paterson and Tinker (2) had found did not differ by a significant amount in reading legibility. They were Bodoni, Caslon Old Style, Cheltenham, Old Style, and Scotch Roman. The sixth printing type used was Kennerly which is also quite legible but was not used in the Paterson and Tinker study. Two syllables were printed in each of the 6 different kinds of type per list, and no nonsense syllable was printed in the same type more than once in the 6 lists. Table 1 shows the order in which the syllables appeared in the various type faces.

5. *Subjects and Grouping.*—Two hundred twenty-seven students (mainly sophomores) in 7 Business Psychology classes and one General Psychology class were used. Each class was subjected to the experiment as a group in its regular classroom at the regular class hour. The 12 nonsense syllables, repeated 12 times, were shown to 126 of the students whose attention was directed toward the fulfillment of other tasks than the learning of the syllables. This group will be called the 'non-intent' group. In the 'intent' group, the students were instructed to learn the syllables. For 72 of them the 12 nonsense syllables were repeated 5 times, and for 29 others, 6 times.

6. *Training of the Non-Intent Group.*—The subjects in this group were given their trials before the intent group so that there would be less likelihood of their obtaining any information about the nature of the experiment. Twelve-page tablets were passed to the subjects comprising the non-intent group and the subjects were asked to write each syllable as it was seen. This use of tablets was followed for both groups so that the experimenters could be sure that the subjects had perceived each syllable accurately. The subjects were told to select the most legible printing type face for each of the 12 lists of 12 nonsense syllables.

The complete instructions given to the non-intent group were as follows:

> Please put your name on the tablet given you.
>
> This is a practical psychological experiment to determine the legibility of different kinds of printing type. The 6 different kinds of type used in this experiment are illustrated on this card.[3] The names of the type are unimportant as far as your part in the experiment is concerned, but are named Bodoni, Caslon, Cheltenham, Kennerly, Old Style, and Caslon Old Style. As you see, the printing is different although the differences are not great. There have been other experiments on the coarser differentiation of type legibility, but this one is to narrow down the finer differences in legibility of these commonly-used types.
>
> The procedure is this: I have a series of nonsense syllables and shall expose each syllable to you for just a thirtieth of a second. You will be shown one syllable and then you are to write immediately on the tablet already provided you, the syllable which you will have just seen. You will do this for all of the syllables used. I will group them by twelves and you are to write only 12 exposures to a page. Since this is a test on legibility, what you are to do is to mark the most legible syllable in each list of 12. It is immaterial whether you are acquainted with the printing types; just mark the one syllable on each page that is easiest to see. It has been found easiest to do this if you put a small check after the syllables that tentatively seem the easiest to read, and then, after the 12 have been presented, draw a big check after the most legible of that list. Now, some of the syllables will be repeated in subsequent lists in different printing types. The purpose of this repetition is to be sure that the differences in legibility are due to the printing type alone, and not to the easier-read shape of some of the letters. This repetition need not concern you as far as working the experiment is concerned. After completing each page, crease it and turn it over. Never turn back nor refer to your neighbor's tablet as a check on your own, but concentrate on the material being presented.
>
> This is a scientific experiment and all possibility of suggestion must be eliminated so as not to affect the results. I shall calculate your ranking of the different types, and I will let you know how you stand in respect to the class as soon as possible. Be attentive, because a thirtieth of a second is not a very long exposure. In each list of 12 syllables, be sure to

[3] On the card was an illustration of each of the various printing types used. The letter 'P' was used for the illustration in each case because it was typical of the variation in types and because this letter was not used in the syllables presented in the lists. This latter was necessary so that additional learning or suggestion would not take place in the non-intent group.

mark the one that is easiest to see. To keep conditions constant, questions will not be answered during the experiment. Are there any now?

The 12 presentations of the syllables were then given to the class. The time between the presentation of successive lists was kept fairly short, just enough time being allowed for the subjects to make the necessary checks and to turn the page of their tablets. The subjects in the intent group also were allowed only this same short interval. Following the 12 presentations, the tablets of the non-intent group were collected and recall sheets quickly given to the subjects and the following instructions read:

> Now write in any order all the syllables you can possibly remember of the 12 presented. We will work out the class average, and let you know where you stand as soon as possible. Please work out your own list, because this is a scientific experiment on learning without the intent to learn, and it must be as accurate as possible. It is important that you remember as many as possible. Be sure to write your name on your list of remembered syllables. Guess if you are not sure. I have one more request; do not tell anyone about this experiment until after lunch today, because other classes must be subjected to the same conditions.

The time allowed for recall was 4 minutes. The scoring of the tablets and recall sheets is described in sections 8 and 9 under Apparatus and Procedure.

As far as the experimenters could ascertain, no one in the non-intent group suspected that he was going to be asked to recall the syllables after their presentation.

7. *Training of the Intent Group.*—Five-page tablets (6-page tablets to those receiving 6 trials) were given to the subjects of the intent group. Seventy-two students were presented 5 of the 6 lists of syllables and told that they were to be tested for retention of them immediately after the presentation of the fifth list. The complete instructions were as follows:

> Please put your name on the tablet given you.
>
> This is a psychological experiment to determine the amount of memorization of nonsense syllables possible when the list of syllables is presented 5 (6 with one group) times. The procedure is this: I have a series of nonsense syllables and shall expose each syllable to you for just a thirtieth of a second. You will be shown one syllable, and then you are to write immediately on the tablet already provided you, the syllable which you will have just seen. Shortly thereafter, a second syllable will be exposed and again you are to write down what you have seen. You will do this for all the syllables used. I will group them by twelves and you are to write only 12 exposures to the page.
>
> Each presentation of the list contains exactly the same 12 nonsense syllables, but each in an entirely different order. After completing each page, crease it and turn it over. Never turn back nor refer to your neighbor's tablet as a check on your own, but concentrate on the material being presented. This is a scientific experiment, and all possibility of suggestion must be eliminated so as not to affect the results. Now this is important to remember: later during the hour I shall ask you to reproduce as many of the syllables as possible. I shall calculate the class average and if you are interested I shall let you know as soon as possible where you stand.
>
> Be attentive, because a thirtieth of a second is not a very long exposure. To keep conditions constant, questions will not be answered during the experiment; are there any now?

Immediately after the presentation of the syllables the tablets were collected and the recall sheets distributed. The instructions read to the subjects then were as follows:

> Be sure to write your name on your list of remembered syllables. Write in any order all the syllables you can possibly remember of the 12 presented. It is important that you remember as many as possible. We will work out the class average and let you know where you stand as soon as possible. Please work out your own list, because this is a scientific experiment on learning, and it must be as accurate as possible. If you are not sure about your syllables, guess.

The time allowed for recall for this group was also 4 minutes.

Twenty-nine additional students were given 6 presentations of the lists. This was necessary because the average recall score of the intent group was slightly below that of the non-intent group. Since it was impossible to expose between 5 and 6 presentations of the lists in order to

equate the two groups, this class of 29 students was given 6 presentations and was included in the intent group to raise its average immediate-recall score. It was important to keep the group averages the same since members of the two groups were also to be matched with each other mainly on the basis of the number of syllables reproduced.

8. *Criteria for Elimination of Unusable Subjects.*—In order for a subject to have had an opportunity to learn the syllables, it was necessary for him to have perceived the syllables correctly. The method used by the experimenters for checking this was to analyze the tablets on which the subjects had recorded what they had seen as each syllable was presented. When an individual recorded a syllable incorrectly there was a great likelihood that the syllable had not been perceived correctly. Such errors are called 'perceptual errors' throughout this article.

Two scores for an individual, based upon his perceptual errors, were considered as being important in determining whether or not that individual's learning and retention data could be used: (*a*) the total number of syllables incorrectly perceived, and (*b*) the number of errors of perception occurring on each syllable. The first score was important to consider in order to eliminate those subjects with poor vision, slow perception time, or other difficulties that interfered with their seeing and recording the syllables correctly. The second measure was necessary in order to eliminate those subjects that misread any one syllable so frequently that the chances of its being recalled correctly were small.

Upon analysis of the error scores for the non-intent group, it was decided that a total error score of 11 perceptual errors (or less) was not only rigid enough to insure accurate perception, but also lenient enough to give a relatively large number of subjects. This total error score allowed for an average of less than one error per page (list) in the 12-page booklet.

Since the non-intent group had 12 presentations of the lists and the intent group had only 5 (or 6) trials, it was necessary to work out a ratio to use as a standard for comparing the groups in terms of the total number of perceptual errors. If, one assumes that the perceptual errors were evenly distributed throughout both sets of tablets, 12 perceptual errors for the non-intent group would be equated to 5 errors for the intent group. However, upon counting the errors on the first 5 pages of the 12-page tablets it was found that 53 percent of the perceptual errors occurred in these first 5 pages. This gave a ratio of 1.89 to 1.00 for the number of perceptual errors in the 12 pages to the first 5 pages of the tablets of the non-intent group. Since the physical conditions of perception for the two groups were comparable, it was assumed that the same ratio of 1.89 to 1.00 would also hold between the total number of errors in the 12 pages in the tablets of the non-intent group and the 5-page tablets of the intent group. Using the ratio of 1.89 to 1.00, this meant that no more than 6 (5.82) errors should be allowed in the 5-presentation intent group. In the case of those individuals in the intent group who received 6 presentations of the lists, their perceptual error score on the first 5 pages of their tablets was used and assumed to be comparable to the error score for the 5-trial group. All individuals in the non-intent group with more than 11 perceptual errors and all those in the intent group with more than 6 perceptual errors (on the first 5 pages for the 6-presentation group) were eliminated as unusable subjects because of too many errors.

If a subject made too many errors of perception on one syllable he obviously would be handicapped in his correct recall of that syllable. In order to prevent errors of this sort from influencing the results, if subjects in the non-intent group made more than 7 errors on any one syllable, or if the subjects in the intent group made more than 3 on any one syllable in the 5 presentations (first 5 presentations in the 6-presentation group), those subjects were considered as unusable and were eliminated. Although the main criterion in matching individuals in the two groups was the immediate recall score, these attempts were also made to equate the original perception.

9. *Matching of the Subjects of the Two Groups.*—Since the number of trials given to each of the two groups was so planned that the average number of syllables recalled immediately after the presentation of the lists was the same for the two groups, it was possible to compare the retention of the groups at some later time when retention tests were to be given. Not only were the total groups to be compared, but a refinement of the comparison was introduced by using the method of matching subjects in the two groups in terms of their scores obtained immediately after the lists had been presented.[9] These subjects were then compared on their later retention scores. In the matching of these individuals it was important that the two total groups of usable subjects averaged the same number of syllables immediately recalled so that in the matching process subjects in the poorer group who were better able to memorize were not matched with the poorer subjects in the better group.

Subjects in one group were matched with subjects in the other group that (*a*) had a comparable number of perceptual errors (an attempt was made to follow the ratio of 1.89 to 1.00[4] mentioned in the previous section), (*b*) had the same number of nonsense syllables correctly reproduced immediately after the presentation of the lists, and (*c*) were present in both the learning and the later recall and recognition situations referred to below. The second criterion was, of course, very important for making the original learning in the two groups comparable. Six pairs of subjects were included in the total group even though not perfectly matched on criterion *b*. The scores in each of these pairs, however, did not differ by more than one syllable; three pairs differed by having the subject in one group superior and three differed by having the subject in the other group superior. Forty-two pairs were perfectly matched for criterion *b*, making a total of 48 pairs of subjects for comparison. The average number of syllables recalled for these two groups was therefore exactly the same, 6.08. The range of scores used in matching was between one syllable recalled and 12 recalled.

A distribution of the reasons for excluding subjects from the final data is shown in Table 2. One hundred seven subjects were not used in calculating the results in this experiment. This elimination of unusable subjects did not seem to have introduced any selective factor in one group more than the other. All of the subjects in the intent group that met the criteria were matched with subjects in the non-intent group, but there was an excess of unmatched subjects in the non-intent group because of the much larger number in this group from the beginning. These unmatched subjects that met the criteria are used, of course, in the total group comparisons.

TABLE 2

REASONS FOR EXCLUDING SUBJECTS FROM THE VARIOUS COMPARISONS

Reason	Group	
	Intent	Non-Intent
Too many errors of perception	42	41
Absent on retest	7	11
"Didn't have glasses," "Couldn't see," or did not write out syllables at time of their presentation	4	2
Total number of subjects above not used in calculating the results	53	54
Number of qualified subjects not included in matched groups because they could not be matched for comparable number of syllables immediately recalled	0	24
Qualified subjects used in matched groups	48	48
Total number of subjects	101	126

TABLE 3

COMPARISONS BETWEEN THE TWO TOTAL GROUPS IN RECALL AND RECOGNITION SCORES

	Total Groups				
	Intent—$N = 48$		Non-Intent—$N = 72$		Critical Ratio
	M	σ_M	M	σ_M	
Immediate recall	6.08	0.36	6.57	0.28	1.07
19-day recall	1.92	0.26	2.08	0.17	0.53
19-day recognition	7.88	0.26	8.47	0.18	1.84

[4] In this matching of subjects the nearest that this ratio could be approximated and a maximum number of subjects still be obtained was 1.76 to 1.00.

10. *Tests for Retention.*—After a 19-day interval, which included a 9-day spring vacation, a previously unannounced retest was given to the students in both groups. These retests were given in the same classroom in which the previous learning had taken place. The experimenter's instructions to the classes were as follows:

> Those of you who did not participate in the first part of this experiment on the learning of nonsense syllables before spring vacation, can just remain quiet and not disturb the rest of the class.
>
> We are interested in determining the number of nonsense syllables it is possible to recall after a 19-day interval, when the original list of nonsense syllables has not been completely learned. On the paper handed you, please write your name, and as many of the 12 nonsense syllables as possible. If you are not sure of the syllables, guess, because many times the guesses are correct. You may possibly recall some nonsense syllables which are different from the ones you recalled in your first list.
>
> Do not refer to your neighbor's list, but work on your own for the few minutes which will be given you to recall the 12 nonsense syllables. Please work out your own list because this is a scientific experiment on retention and must be as accurate as possible.

Four minutes were allowed for reproduction of the syllables retained. After these slips were collected, a mimeographed list of 60 nonsense syllables from Glaze's list (1) of syllables of zero percent association-value, including the 12 syllables presented 19 days previously, was then passed to each subject. The following instructions then were read to the subjects.

> You may have been under too great a handicap to reproduce the nonsense syllables after such a long time, so I shall let you attempt to recognize the original 12 which are included in a list of 60 nonsense syllables. Be sure to check exactly 12, no more nor less, even though you have to guess. I should also like you to put a second check beside those of which you are quite certain. Write your name on this sheet. Do not refer to your neighbor's sheet, but help keep this as objective and scientific as possible.

Four minutes were allowed for recognition and then these papers were collected.

Results

The two total groups of usable subjects are first compared for the number of syllables recalled immediately after the presentation of the lists in order to find whether or not the two groups had learned approximately the same number of syllables with a different number of presentations of the lists. Distributions showing the number of syllables reproduced immediately after the learning by the two groups are shown in Fig. 2. Since the *N*'s for the two groups are different (intent group = 48; non-intent group = 72), the two curves are plotted in terms of the percent of subjects that recalled the various numbers of syllables. The distribution curves for the two groups are seen to be quite comparable, although the mean of 6.57 syllables reproduced for the non-intent group is slightly higher than that of 6.08 for the intent group. The critical ratio for this difference is 1.07, and is interpreted as indicating that the difference obtained between the means of the immediate recall scores for the two groups is not statistically significant. Table 3 shows these averages as well as the other statistical comparisons between the two total groups.

The 19-day recall scores for the two total total groups of usable subjects that were present for the retest, average 1.92 for the intent group and 2.08 for the non-intent group. The critical ratio for this differ-

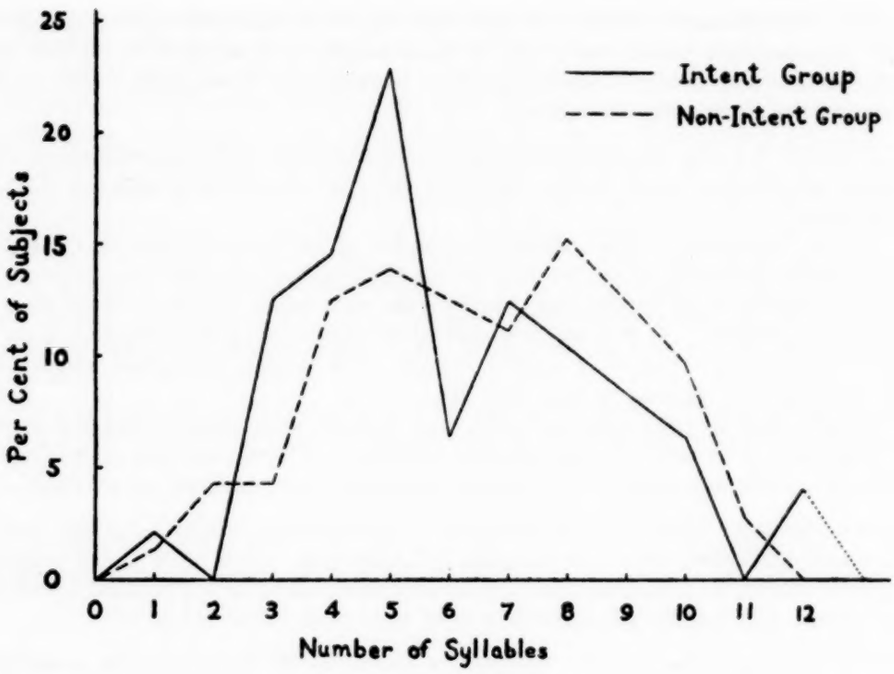


FIG. 2. Percent of subjects in each group who recalled the different numbers of nonsense syllables on immediate recall.

ence is 0.53, which indicates that the obtained difference is not statistically significant. As far as the total group comparison is concerned, material which is learned equally well either with or without the intent to learn is retained equally well. This interpretation also holds for the mean recognition scores, which are 7.88 syllables recalled for the intent group and 8.47 for the non-intent group. The critical ratio is 1.84. The non-intent group which was slightly ahead in the original learning is slightly superior throughout these and later comparisons.

The main comparisons here, however, are between the two groups of 48 subjects which had been matched in terms of their immediate recall scores. The data for the comparisons between these groups are shown in Table 4. The average recall score after the 19-day

TABLE 4

COMPARISON BETWEEN THE TWO MATCHED GROUPS IN RECALL AND RECOGNITION SCORES

	Matched Groups				
	Intent—N = 48		Non-Intent—N = 48		Critical Ratio
	M	σ_M	M	σ_M	
Immediate recall.....	6.08	0.36	6.08	0.35	0.00
19-day recall.....	1.92	0.26	2.08	0.22	0.70
19-day recognition.....	7.88	0.26	8.52	0.22	2.21

interval for the intent group is 1.92 and for the non-intent group, 2.08. In calculating the standard error of the difference between these two means, the correlation of +.54 between the two groups of matched subjects on their 19-day recall scores was, of course, taken into consideration. The critical ratio for this difference is 0.70 which indicates that the difference is not statistically significant. These results indicate that the retention of nonsense syllables is approximately the same whether they are learned with or without the intent to learn, under the conditions of this experiment.

In the case of the retention scores of the matched groups when measured by recognition, the mean number of syllables recognized by the intent group is 7.81 and by the non-intent group, 8.54. The correlation between these two sets of scores is +.33. The critical ratio for the difference between the means when this correlation is taken into consideration is 2.21. Although this figure approaches significance, the fact that the critical ratio for the comparison between the recall scores in the matched groups is small causes little weight to be attached to this non-significant critical ratio. These results are interpreted as indicating that no statistically significant differences exist between the groups.

Discussion

The results obtained in this experiment indicate quite clearly that when nonsense syllables have been learned to the same level of mastery by groups learning under conditions of intentional and of incidental learning, the material learned tends to be retained equally well after a 19-day interval. These findings do not agree with the findings in a majority of the previous studies on this problem because in the previous studies the two groups of subjects originally had not learned the material equally well. As a result, the group which had learned the material less well (incidental group) appeared to have retained less of the material.

In the present experiment the directions given to the non-intent group during the original observation of the nonsense syllables undoubtedly caused the physiological traces of the experiences to be organized differently than were the traces in the intent group. When the directions were given then to the non-intent group to recall the syllables, a reorganization of the traces was necessary. The subjects in the intent group, of course, had been learning, or attempting to learn, the material from the first trial. The trace organization resulting from their attitude while learning made it possible for them to learn the material as well as the non-intent group in fewer trials.

Of course, with the material being of a nonsense character, and with there being no repeated contiguity between the syllables, organization in both groups was difficult. Through our preliminary experimentation before beginning the present experiment, we had by trial and error found a balance of trials between the two groups which would produce approximately the same amount of organization and immediate recall. Apparently in the non-intent group the amount of trace organization which resulted from the change in directions was just equal to the amount of organization which had developed in the intent group.

After the immediate recall of the nonsense syllables in both groups, the retention of the material undoubtedly was considered as a completed task and rapid forgetting began. Since the forgetting is brought about not only through the fading of these traces with time, but also by the interference with these traces by succeeding activities, this fading, or interference took place in both groups with approximately equal speed regardless of which of the two conditions of learning preceded the recall.

Summary

This experiment was designed to compare the effects of intentional and incidental learning on the retention of nonsense syllables when the amounts of original learning were equal.

Twelve nonsense syllables of zero association-value were presented in a planned, random order to two groups of subjects by means of a tachistoscope for .03 second each. Six easily legible but slightly different types of print appeared in each presentation of the list. The usable subjects in the incidental group ($N = 72$) were instructed to observe the syllables and to indicate the most legible printing type for each of the 12 presentations of the list. The usable subjects in the intent group ($N = 48$) were instructed to disregard the printing-type differences and to learn the list of 12 nonsense syllables which was presented to them 5 (or 6) times. Average immediate-recall scores for the two groups were approximately the same.

Forty-eight subjects in the non-intent group were matched with the 48 in the intent group on the basis of the number of errors made during observation of the syllables and the number of syllables reproduced correctly after presentation of the material. Recall and recognition scores after a 19-day interval showed no statistically significant differences either between the two total groups or the two matched groups.

These results indicate that nonsense syllables which have been learned to the same level of mastery are retained equally well if

learned originally either under conditions of incidental learning or of intentional learning.

(Manuscript received August 3, 1942)

References

1. Glaze, J. A. The association value of non-sense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
2. Paterson, D. G., & Tinker, M. A. Studies of typographical factors influencing speed of reading. X. Style of type face. *J. appl. Psychol.*, 1932, **16**, 605–613.
3. Peterson, J. Effect of attitude on immediate and delayed responses; a class experiment. *J. educ. Psychol.*, 1916, **7**, 523–532.
4. Thisted, M. N., & Remmers, H. H. The effect of temporal set on learning. *J. appl. Psychol.*, 1932, **16**, 257–268.

AN EXPERIMENTAL STUDY OF 'REPRESSION' WITH SPECIAL REFERENCE TO NEED–PERSISTIVE AND EGO–DEFENSIVE REACTIONS TO FRUSTRATION

BY SAUL ROSENZWEIG

Worcester State Hospital and Clark University

Elsewhere (1) a theoretical distinction has been drawn between need-persistive and ego-defensive reactions to frustration. The former type of response concerns the fate of the frustrated segmental need, the latter the fate of the individual as a whole in a situation of stress. Whereas need-persistive reactions serve to fulfill the frustrated need despite momentary obstructions, the ego-defensive ones tend to protect the integration of the personality if and when the latter is threatened with disruption. Most behavior incident to frustration involves both types of reaction but pure cases of each alone are found empirically and the distinction has theoretical advantages besides.

The relation of the present distinction to current psychoanalytic concepts is obvious. One function of the formulation is, in fact, to clarify the difference among psychoanalytic mechanisms and to assist in their experimental study. Thus, the concept of sublimation points chiefly to need-persistive reactions; projection concerns the defense of the ego. A fusion of need-persistive and ego-defensive reactions is found in the mechanism of reaction-formation since here an inhibited impulse gains indirect satisfaction (sublimation) while protecting the ego by a kind of displacement (projection). Nearly all of the psychoanalytic mechanisms could be similarly allocated in the present framework.

If the attempt is made to characterize the mechanism of repression in this setting, it at once becomes clear how complicated repression really is. What can readily be seen is that, like reaction-formation, repression involves both the inhibition of some need—in classic instances, the sexual one—and at the same time the defense of the ego. Even at first glance, therefore, both need-persistive and ego-defensive reactions seem to play a part.

As a demonstration of the distinction between need-persistive and ego-defensive reactions to frustration and as a contribution to the experimental psychology of repression, a study conducted on a group of 60 college students at the Harvard Psychological Clinic several years ago (2) may be reported. In an earlier investigation (3) of 40

children ranging in age from 4 to 14 years, it was found that one group tended to recall finished (successful) tasks more often than unfinished (unsuccessful) ones, while another group did the opposite. Ratings for the trait of pride assigned by the children's teachers were, on the average, higher for the former than for the latter group. The experiment to be reported was designed in part as a verification of these results. In it an attempt was made to control the external experimental situation so as to reproduce with two groups of 30 adults each the alternative conditions which had obtained with the children by virtue of certain age-linked differences in their developing personalities.

Procedure

Each subject was given a series of jigsaw puzzles to solve but was deliberately permitted to finish only half of them. The puzzles represented such common objects as a boat, flag, house, grapes, etc., and were pictured in miniature completed form on small cards which the subject was permitted to study for approximately a quater of a minute before being given the corresponding task. In 40 of the cases, 20 in each of the two groups, the same puzzles were presented in the same order with the same outcome for the subjects in both groups. In the remaining cases, the puzzles were more varied and presented in more random order but, with a few negligible exceptions, each task finished or unfinished in one group was matched by a corresponding outcome for that task in the other group. An attempt was also made to equalize the amount of time spent on the successful and the unsuccessful tasks. This objective was attained by interrupting a subject on some puzzle at the end of an amount of time equal to what he had spent on some previously finished task.

Recall was requested by the experimenter immediately, *i.e.*, within one minute, after the work on the last puzzle was concluded. Handing the subject a sheet of paper and a pencil, the experimenter said: "Will you now please write down in any order in which they come to mind the names of all the puzzles I asked you to do. Tell me when you can't remember any more." The recall was terminated by the subject, usually by his handing back the sheet of paper, and in nearly every case was completed within five minutes.

The difference in the conditions for the two groups of subjects, which constituted the point of the experiment, centered around the arousal of need-persistive reactions alone or of ego-defensive reactions in addition. The subjects in one of the two groups—to be designated the *informal* group—were obtained from the Student Employment Office at a small fee per hour. The experimenter informed them that he had constructed certain puzzles to be used in some experiments in the near future but that it was necessary to determine how they worked out in actual practice. Thus, the subjects were told, the present session would aid the experimenter in finding out what he wanted to *know about the puzzles*. To this end the work on any task might be interrupted as soon as the experimenter got the necessary information. Such interruption would obviously not reflect at all upon *the ability of the subject*, it was implied, since, if anything, the sooner the information desired was obtained, the more successful the session would be. The exact protocol of instructions for this group was as follows:

> (These instructions are not to be read but are to be repeated from memory in a conversational way.)
>
> Have you ever done jigsaw puzzles?.....
>
> Do you like doing that sort of thing?.....
>
> .....Well, I'm going to ask you to help me by doing some jigsaw puzzles for your amusement and my edification. That is to say, this is not in any way a test of your ability or anything else about you. The point is simply this: I intend to use these puzzles in certain experiments later on. Before I do so, however, I wish to get some idea of how they work out in actual practice and so I am doing this preliminary work with you. Is that clear?
>
> You will find that every puzzle comes out as a board one foot square. Also every puzzle has a picture represented on it and I shall let you study a small card with a miniature picture of the completed puzzle on it for a little while before giving you the pieces for each puzzle.

Don't hurry or feel in any way constrained while doing these puzzles as that would spoil the whole purpose of the thing. I'm not in the least interested in seeing how well or how quickly you can do puzzles. Moreover, do not be surprised if I interrupt you before you finish some of the puzzles for I doubtless shall do so if I find out what I want to know about a particular puzzle before you finish.

Have you any questions?.....
.....Then we can begin.....
.....(Then present picture.).....
.....Here is a picture of a..... Look at this for a little while.....
.....All right?.....
.....Here are the pieces for this puzzle.

It will be seen that the subjects in this group were expected to be working under conditions of comparatively little tension. The extent to which these expectations were satisfied may be gathered from the introspective reports. For example, one of the subjects wrote: "Very calm; mildly interested. . . . Not a very important matter. No hurry. Plenty of time (most pronounced feeling). It occurred to me that fear was entirely absent."

The individuals in the other group—which may be designated the *formal* group—were enlisted from the Freshman student advisees responsible to the Director of the Clinic and were personally invited by him to participate. They were given the tasks to perform as an intelligence test. In contrast to the informal group, every effort was here made to arouse tension by such expedients as holding a stop-watch visibly in front of the subjects and by presenting instructions succinctly and curtly. Not the puzzles but the subjects were being tested in this case. Interruption was to be interpreted not as a favorable result but as failure due to the subject's shortcomings. The exact protocol of instructions for this group was as follows:

Have you ever done jigsaw puzzles?.....
Do you like doing that sort of thing?.....
.....The point is that I am going to give you some jigsaw puzzles to do as an intelligence test so that you may be compared with the other persons taking the test.....
.....Now in order that you may get some conception of what you are expected to do, I shall briefly describe the procedure to you.

Each puzzle has a picture represented on it. I shall begin in each case by showing you a small card which has on it a miniature of the completed puzzle. You are to study this for about a quarter of a minute (18 sec.) and you are advised to do this carefully for you will not be allowed to look at the picture again. The completed puzzle will in every case be in the form of a board one foot square; note that this is somewhat smaller than the board in front of you on which you will work. Is that clear?

Now each puzzle has one and only one possible solution and you are allowed a specified time in which to find it. Every puzzle counts as much as every other in the score. However, different times are allotted the different puzzles according to their difficulty. If you do not solve any puzzle in the allotted time, I shall naturally be obliged to stop you. You ought therefore to work as rapidly as you can for the test is so scored that you get either full credit or no credit for each puzzle. Your work will be interpreted as representing the full extent of your ability, so do your best.

Are there any questions?.....
.....If not, we are ready to begin.....
.....Here is a picture of a..... Study this for about a quarter of a minute (18 sec.).....
.....Time.....
.....Here are the pieces. Please wait each time till I remove all the pieces from the box and say 'Go.'
.....Go.

The extent to which the instructions succeeded may again be gleaned from the introspective reports. A sample is as follows: "My feeling was one of desperation. With each puzzle that I missed the thought forced itself upon me that I wasn't getting on well, that I had no business being at Harvard or in the field of Philosophy." Another subject stated: "I felt very impatient when I couldn't solve some of them and wanted to push them away from me when I couldn't finish in time."

It must, however, be pointed out that the instructions occasionally miscarried. This sort of result was less characteristic of the formal group than of the informal one. An example from one subject in the latter group may be given in these words from his introspective report: "I had very few feelings other than a certain tense desire to solve the puzzles in as quick a manner as possible with as few movements. . . . In addition I was conscious that I was being watched, however honestly I tried to prescind from that. There was a great desire not to be overcome, to make a good showing." Such exceptional attitudes on the part of the subjects in the two groups must naturally be interpreted as indicating that occasionally, if not frequently, the personality traits of a subject may override the environmentally imposed conditions of an experiment. In such instances results running counter to a hypothesis will, of course, be taken as inconclusive. In the present experiment, since comparison of groups, with 30 subjects in each, was in question, individual exceptions like the foregoing would not be expected to counteract any general difference which the experimental conditions were calculated to create.

From the technical standpoint it should further be observed that not only were the same tasks in general given the subjects in both groups but that the time allowances for performance on the several puzzles were about equal for both. The fulfillment of this latter condition in the informal group was aided and made available for measurement by the coöperation of an assistant who sat at one side of the experimental room, presumably occupied with some unrelated task of her own. At the beginning of the session the experimenter asked her permission to work with the subject in the room and she graciously explained that she was not doing anything which would be seriously disturbed. As a matter of fact, this assistant had in front of her a Telechron clock from which she could read off the time spent on each task given the subject. Since the picture on the task was announced clearly as the corresponding puzzle was given the subject, the assistant could easily keep a record of the names of the puzzles and of the time spent on each. She was also able to make certain additional behavioral notes.

Results and Discussion

It will be readily appreciated that the conditions for the two groups of subjects differed primarily in the extent to which tension was evoked. Perhaps, more accurately stated, the proposition would be that any tension aroused with the informal group of subjects would pertain primarily to the accomplishment of the task in hand *qua* task. On the other hand, the subjects in the formal group were intended not only to have such task-tension aroused in them but were in addition being aroused in a personal way that would involve their attitudes of self-esteem, pride, etc. That is to say, the subjects in the informal group would in general have no occasion to regard the incompletion of a task as an experience of failure on their part. If anything, interruption by the experimenter at an early point might be construed as showing that the subject had revealed the desired information with unusual alacrity—a successful result. By contrast, subjects in the formal group were personally involved in the performance of the tasks. For them interruption was equivalent to failure. At the same time these subjects might also be supposed to have had certain residuals of tension related to the tasks as such. The hypothesis of the experiment in accordance with these circumstances was that under the informal conditions, the unfinished tasks would be better recalled than the finished ones because need-persistive responses alone would be operative and would make for the easier recall of tasks with which undischarged tension was associated. Such a result could be ex-

pected not only theoretically but on the basis of previous work done by Zeigarnik (4). Conversely, subjects in the formal group were expected to recall finished tasks—in this case successful ones—more frequently than unfinished or unsuccessful ones, the assumption being that with the arousal of pride and accompanying ego-defense in case of failure, the individual's needs for inviolacy would take precedence over the task-tension making for recall of the unfinished tasks. As a matter of fact, the findings of the experiment substantiate the hypothesis, with certain qualifications to be noted presently. Table 1 shows the experimental results individual by individual. Table 2 summarizes them for the groups.

In the analysis of these results the individual subject will be taken as the unit and Table 2 will provide the chief data. The justification

TABLE 1

MEMORY RESULTS FOR INDIVIDUALS WITH CONDITIONS OF NEED-PERSISTENCE AND EGO-DEFENSE

Subjects with Informal (Need-Persistive) Conditions					Subjects with Formal (Ego-Defensive) Conditions				
Name	F* Given	U Given	F Recalled	U Recalled	Name	F Given	U Given	F Recalled	U Recalled
Bea	9	9	6	2	Ban	9	9	3	6
Ben	9	9	5	7	Bar	9	9	4	6
Ch	9	9	3	6	Br	9	9	7	6
Cl	9	9	4	6	Camo	9	9	3	2
Cra	9	9	7	7	Camp	9	9	8	8
Cre	9	9	3	5	Da	9	9	7	7
En	9	9	4	5	Dr	9	9	4	3
Fi	9	9	4	5	Fi	9	9	5	4
Fr	9	9	6	5	Gr	9	9	8	7
Gi	9	9	6	8	Ha	11	11	6	9
Ha	9	9	8	8	Ho	9	9	6	6
Hu	9	9	5	6	Ka	9	9	4	6
Ka	9	9	5	7	Le	9	9	4	5
Ke	9	9	9	4	Mc	9	9	5	4
La	9	9	6	6	Ni	9	9	6	5
Le	9	9	1	5	Pa	9	9	7	7
Ly	9	9	6	6	Por	9	9	5	3
Mc	9	9	4	7	Pow	9	9	7	8
Men	9	9	2	5	Smi	9	9	7	2
Mer	9	9	6	7	Smy	9	9	6	4
Mo	9	9	5	2	Sn	9	9	5	6
Ola	9	9	5	7	Ta	9	9	7	5
Olk	9	9	5	4	Trol	9	9	6	6
Po	5	5	5	3	Tros	9	9	5	6
Re	9	9	4	5	Wal	9	9	5	4
Sa	9	9	4	6	War	10	10	5	4
Un	9	9	7	6	We	9	9	5	4
We	9	9	3	4	Wh	9	9	8	7
Ya	9	9	3	6	Wi	9	9	5	9
Yo	9	9	5	7	Za	9	9	6	5
Totals	266	266	146	167		273	273	169	164

* F signifies finished, U unfinished tasks.

TABLE 2

MEMORY RESULTS FOR GROUPS WITH CONDITIONS OF NEED-PERSISTENCE AND EGO-DEFENSE

	Number of Subjects Who Recalled Preponderance of Finished Tasks	Number of Subjects Who Recalled Preponderance of Unfinished Tasks	Number of Subjects with No Preponderant Tendency
Group with informal conditions	7	19	4
Group with formal conditions	17	8	5

for so doing is that from the standpoint of the psychological significance of the behavior involved, the subject *is* the unit. The comparison of the experiment is not between *task* and *task* but between *subject* and *subject* working under *different* conditions. To ignore the individual subjects and consider the tasks as tasks, or to pool the results in terms of remembered tasks for all subjects would falsify the findings. This consideration is especially important because the differences for each subject in the recall of finished and unfinished tasks were not large. They ranged from a preponderance of 1 to a preponderance of 5, with approximately 18 puzzles, 9 finished and 9 unfinished, worked by each subject. (*Cf.* Table 1.) The results are therefore presented as for two contrasted *groups of subjects* and the conclusion is construed as demonstrating only a tendency or a difference in tendencies under two contrasted sets of conditions.

It will readily be seen from Table 2 that, subject for subject, memory in the informal group favored the unfinished tasks 19 to 7. (Four of the subjects showed no preponderant tendency.) Here, apparently, the need-persistence related to the incomplete tasks enhanced the chances of their recall. With the formal group, where pride had been aroused and ego-defense activated by failure, the memory results favored the finished or successful tasks 17 to 8. (Five of the subjects in this group showed no preponderant tendency.) The chi square test of these values gives a figure of 8.76 with a P value of $<.02>.01$. If the distributions are tested with the subjects who showed no preponderant tendency omitted, chi square is found to be 8.64, with a P value of $<.01$. The conclusion thus seems statistically warranted that the subjects in the two groups produced results which are significantly different—a result which would be expected to occur by chance less than once in a hundred times.

It may be pointed out in passing that in view of the fact that the subjects in the formal group may be supposed to have had a certain amount of need-tension associated with the tasks as such, the present findings are perhaps even more significant than appears. Under the experimental conditions, the subjects in the formal group were presumably being influenced by task-tension to recall the unsuccessful

tasks and their ego-defenses had to counteract or work beyond this tendency before successful tasks, *i.e.*, finished ones, were better recalled. This interpretation is borne out by the differences in the relative preponderance scores for the two groups. If for the purpose of such a comparison the index [1]

$$100\left(\frac{\text{finished} - \text{unfinished tasks recalled}}{\text{finished} + \text{unfinished tasks recalled}}\right)$$

is taken to represent the percentage dominance of finished tasks in the recall of each subject, the algebraic sum of these indices for the individuals in the informal group is -229.6, the mean -7.65; while for the formal group the sum is $+88.4$, the mean $+2.95$. The indication thus is that the tendency for recalling unfinished tasks by subjects in the informal group was considerably more marked than was the corresponding tendency for those in the formal group to recall finished tasks. It may be concluded that if it could have been arranged so that failure was not necessarily coincident with incompletion in the formal group, the difference in the memory results for the two groups might have been even more striking.[2]

In this same context the monopoly which the discussion has up to now tacitly accorded to repression in the defense of the ego must be called into question. As a matter of fact, it is to be expected, on at least clinical and theoretical grounds (**5**), that there exist modes of ego-defense which differ from the sort of impunitive repression here considered in that they aggressively, *i.e.*, intropunitively or extrapunitively, defend the ego by such mechanisms as displacement, isolation and projection. These mechanisms, however, would not necessarily involve the forgetting of unpleasant experiences; they might actually involve some rumination over them at the expense of the more successful experiences. If, in the present experiment, such

[1] The formula here employed is readily generalized, if need be, to include cases where an unequal number of finished and unfinished tasks were performed. The *percentage* of the tasks of a given class, rather than the absolute number of tasks, recalled is then used. The yielded indices range from $+1$, signifying exclusive preference for finished tasks, to -1, meaning exclusive preference for unfinished ones, with zero standing for neutrality. This scoring method avoids certain obvious pitfalls of the one employed by Zeigarnik (**4**).

[2] Another way of making the same comparison, but one which is less representative of the individual results, is to take the total number of finished and unfinished tasks recalled by each group. If this is done, it is found that in the informal group approximately 55 percent of all finished tasks and 63 percent of all unfinished tasks undertaken were recalled; while in the formal group 62 percent of all finished tasks and 60 percent of all unfinished tasks undertaken were recalled. This means that while very nearly the same number of tasks were recalled under both informal and formal conditions (59 percent as compared to 61 percent), the loss in the recall of unfinished tasks by the formal group was less than the gain in the recall of finished ones (3 percent as compared to 7 percent). The interpretation may thus tentatively be made that the subjects in the formal group were still recalling the unfinished tasks by virtue of a fairly strong need-persistive tendency and that their preponderance of finished tasks in recall had to manifest itself over and above this tendency.

additional defenses arose during the formal sessions, one could not expect the predominance of successes in recall—the alleged effect of repression as an ego-defense—to have been much more marked in the results than appears. Indeed, the only explanation that could be adduced for such predominant effects of repression as were observed would need to invoke a presumed greater availability of repression than of other defense mechanisms in subjects and situations like the present.

As regards the predominance of successes over failures in recall during the formal sessions, it is thus clear that this effect could have appeared only by overshadowing the countervalent effects of need-persistence and of certain aggressive types of ego-defense. The fact that such a predominance was actually found would seem to lend considerable support to the concept of repression as a very general mechanism of defense.

These considerations, however, tend to raise a doubt as to the significance of the present findings for the concept of repression itself. Theoretically repression should entail the conscious *forgetting* (and unconscious remembering) of unfinished, *i.e.*, failed and unpleasant tasks; actually the results demonstrate that under the formal conditions, where ego-defense might be expected, the greater recall of successes is more striking than the forgetting of failures. This fact is doubtless attributable in part to the just mentioned countervalent effect of need-persistence related to the unfinished (failed) tasks; the shortcomings of the technique in identifying failure with incompletion is again clear. In partial support of the present findings it can only be repeated that the psychologically sound approach requires the proportion of finished and unfinished tasks in the recall of each subject to be emphasized, not the absolute number of tasks of the two kinds recalled by the subjects collectively.

There is, moreover, a point of view from which the above noted defect of technique may instead be regarded as a virtue. As has been previously stated, repression theoretically involves not merely the defense of the ego from unpleasant memories but also the inhibition of certain interrupted activities. Such inhibited behavior may be conceived as persisting toward fulfillment, directly or indirectly, despite the existence of barriers. It is from such unsatisfied drives that, according to psychoanalytic theory, the conversion aspects of hysteria are energized at the same time that the ego defends itself by forgetting from the unpleasant images associated with the inhibiting trauma. If, now, in the present experiment, it is found that the forgetting of the unpleasant experiences is complicated by a tendency for them to persist in memory because of their being incomplete, a closer approximation to the full concept of repression seems to be

achieved than would be true if completed unpleasant tasks were in question. From this standpoint even the results of the informal sessions, where ego-defense was not intended to figure, would be relevant to the theory of repression and would serve to indicate that inhibited activities persist though 'repressed.' When such inhibited activities are at the same time unpleasant, there is a tendency for them to be consciously forgotten by virtue of ego-defense, but they tend nevertheless toward fulfillment, though perhaps in part unconsciously, through need-persistence. It is thus more readily understandable why the above reported ratios for finished as compared to unfinished tasks in recall indicate a more marked tendency for the unfinished tasks to dominate under the informal conditions than for the finished tasks to dominate under the formal ones. In other words, repression may be regarded as encompassing not only the forgetting of the unpleasant but the (conscious or unconscious) persistence of the unfinished. These remarks are obviously speculative but are offered as pointing the way to further and less superficial work in which methods of probing the unconscious aspects of need-persistive and ego-defensive reactions would be prominently included.[3]

From the standpoint of the ego-defensive aspects of repression, another interesting problem is suggested by the views regarding memory in Breuer and Freud's *Studies in Hysteria* (**6**, p. 7). The difference in conscious memory between the two groups of subjects may have been due, as in their speculations, either to a difference in the degree of *pride* aroused by the experimental conditions or to a difference in the degree of *excitement* evoked. From the present results it is impossible to decide between these alternatives but the question as to whether excitement with its accompanying structural disorganization was associated with the pride related to ego-defense is, like the preceding one, raised here because of its great importance for future experimentation.

One final matter that deserves explicit mention in the present context relates to the individual differences in reaction to frustration which the basic distinction being considered does not in the first instance seem to envisage but which the experimental results bring out saliently. As already noted, subjects frequently superimpose their own personalities onto the environmental conditions which they are expected to observe. Moreover, even when environmental conditions 'take,' types of reaction to frustration may vary widely. Ego-defensive reactions are by no means limited to repression or to the forgetting of unpleasant experiences. Such an impunitive type of

[3] A former discussion (5) of the distinctions which must be made between the conscious and unconscious fates of repressed experiences and between stimulus and response repression is relevant here but differs in approach.

response covers only one possibility. It is, for example, equally possible for subjects to defend their egos extrapunitively or intropunitively. /It must further be remembered that individuals differ not only in their qualitative types of reaction but in the degree to which they react inadequately under the stress of frustration. At this point the concept of frustration tolerance demands consideration.

It seems reasonable to conclude that while the present experiment does not by any means exhaust the possibilities of measuring need-persistence and ego-defense as involved in repression, it at least demonstrates the basic distinction between the two reaction types. As regards the theory of repression, the study brings out the need for critically analyzing this complex concept into its component aspects and indicates one experimental means to this end.

Summary

As a demonstration of the distinction between need-persistive and ego-defensive reactions to frustration and as an experimental study of repression in this setting, an investigation of 60 college students with respect to the recall of finished and unfinished tasks was reported.

Two groups of subjects were given a series of jigsaw picture puzzles to solve. One of the groups was presented the tasks informally with the avowed purpose of helping the experimenter classify the puzzles for future use. The other group performed the same tasks formally as an 'intelligence test.' All subjects were permitted to finish half of the puzzles but were interrupted in each of the remaining half. They were then requested to name the problems which they had attempted. The results were that those individuals who had performed under the informal conditions, where need-persistive responses were alone presumably entailed, recalled the unfinished tasks better than the finished ones (19 remembered more unfinished than finished tasks, 7 did the opposite, while 4 showed no preponderant tendency). Undischarged tension associated with the incomplete tasks was presumably the basis for this effect. The individuals in the 'intelligence test' group, who were presumably motivated in the direction of ego-defense as well as need-persistence, recalled finished tasks more frequently than unfinished ones (17 remembered more finished than unfinished tasks, 8 did the opposite, while 5 showed no preponderant tendency). Here ego-defensive repression may, according to one interpretation, have been instrumental in the forgetting; *i.e.*, the subject's need for inviolacy may have taken precedence over the task-tension making for the recall of the incomplete activities. Individual differences were of course found and are discussed.

It is possible to speculate for purposes of future investigation that the present results indicate the complete mechanism of repression

to involve not only the ego-defensive forgetting of the unpleasant, but the need-persistive retention of the unfinished. Such a view of repression closely approximates the clinical psychoanalytic usage. As regards ego-defensive forgetting, the problem of the interrelated roles of pride and excitement demands further study.

(Manuscript received August 7, 1942)

References

1. Rosenzweig, S. Need-persistive and ego-defensive reactions to frustration as demonstrated by an experiment on repression. *Psychol. Rev.*, 1941, **48**, 347–349.
2. Rosenzweig, S. The recall of finished and unfinished tasks as affected by the purpose with which they were performed. (Abstract) *Psychol. Bull.*, 1933, **30**, 698.
3. Rosenzweig, S., & Mason, G. A. An experimental study of memory in relation to the theory of repression. *Brit. J. Psychol.*, 1934, **24**, 247–265.
4. Zeigarnik, B. Untersuchungen zur Handlungs- und Affektpsychologie, Herausgegeben von K. Lewin. III. Das Behalten erledigter und unerledigter Handlungen. *Psychol. Forsch.*, 1927, **9**, 1–85.
5. Rosenzweig, S. The experimental study of repression. *In* H. A. Murray *et al. Explorations in personality.* New York: Oxford Univ. Press, 1938, pp. 472–490.
6. Breuer, J. & Freud, S. *Studies in hysteria.* Trans. by A. A. Brill. New York: Nerv. and Ment. Dis. Pub. Co., 1936. Pp. ix + 241.

MEASUREMENT OF A MOTOR SET

BY DOROTHY RETHLINGSHAFER

University of North Carolina

Cathcart and Dawson (**2**) found that subjects working with an ergometer tended to have their natural rate of movement displaced toward any intervening rate. Though they stated that individual differences of some consistency were found, they did not further investigate this problem. It was the purpose of this study to measure such a tendency-to-continue a rate of movement by determining the effect of slow or fast rates on the natural tapping rate. In so doing, it is assumed that this is an indirect measurement of the resistance of the muscular adjustment established by the slow or fast rhythms. To the extent that the natural rhythm cannot immediately be restored by the subject, to that extent there is a continuance of the muscular adjustment for the unusual rhythms.

The natural tapping rate can be reproduced with a high reliability,[1] consequently an immediate change in the natural rhythm following the forced rhythm can safely be attributed to the introduced experimental variable. Three specific conditions were investigated:

I. Effect upon the natural tapping rate when subjects were forced to tap at a slower speed than their natural rate, *i.e.*, at one-half their natural rate.

II-a. Effect upon the natural rate when the intervening rate was the same for all subjects, in the first case a rhythm one-half the rate of the *average* of the preceding group. Some subjects were thereby forced into an intervening rate slower than their original natural rate; others were forced to go faster.

II-b. Again, the effect upon the natural rate when the intervening rate was the same for all subjects, but in this case only one tap per second. This forced rate slowed down all but one of the subjects from their original natural rate.

III. Effect upon the natural rate when all the subjects were forced to go faster than their own rate.

[1] Jasper (**3**) found a reliability of .98 on retest of tapping rate following a five minute interval. Allport (**1**) reports corrected reliability of .96 for four measures of hand and finger tapping taken at different experimental sessions. In this study, with varying intervals from three to 15 minutes between the three tests of natural tapping, the estimated reliability was .95.

Subjects, Apparatus, and Procedure

The subjects were 116 students in the general psychology courses at the University of North Carolina. Sixty-one of the subjects were used under the first condition. The remainder were used under the other conditions.

Apparatus.—A Stoelting tapping board was wired in circuit with a Cenco electric counter. A metronome was set according to the rate that the experimenter desired to force upon the subjects. Under Condition II-b, the forced rhythm was obtained by having a Harvard clock activate a telegraph key at one second intervals.

Procedure.—Three records were obtained from each subject of his natural tapping rate by saying: "Tap at a rate that is convenient, natural for you." If any questions were further asked as to whether the rate was to be regular, etc., the directions were repeated. With a few subjects who seemed uncertain how to procede, the further statement was made: "Tap with this stylus on the plate as if you were tapping with your pencil on the table." The time for the subjects to tap 100 times, as recorded by the counter, was taken with a stop watch.

Under Condition I, 61 subjects were forced to tap for one minute to a metronome set at one-half of the average of the three measurements of their own natural rate. Immediately following the forced rhythm the subjects were asked to choose again a rate that was 'natural, convenient' for them.[2] The time was recorded to complete 100 taps.

Under Condition II-a, 55 subjects were asked to tap for one minute to a metronome so set that it took 68 sec. to complete 100 taps. This was double the average natural rate of the preceding 61 subjects. Under Condition II-b, the same 55 subjects were forced to tap for two min. to a telegraph key sounded by a Harvard clock once per sec. This, of course, meant that it took 100 sec. to complete 100 taps.

Under Condition III, 22 of the 55 subjects were asked to tap as rapidly as possible for 10 sec. The 10 sec. period was chosen after experimenting with 30 and 15 sec. as probably sufficiently long to establish a motor continuance effect without creating fatigue. The fast rates adopted by the subjects varied from 80 to 50 taps during 10 sec., with a mean of 64.

Under Condition I the subjects were occupied between trials by various measures of perseveration with which the tapping rates were later compared in another study (4). In Conditions II and III, during the 5 to 6 min. intervals between trials, the subjects worked on the Maier Art Judgment Test.

Results

Various changes in the natural rates might be expected following the forced rhythms: (1) The subjects might be unaffected by the rhythms forced on them and hence choose the same natural rate that they had previously used in three trials with a reliability of .95. (2) The subjects might be so affected by the preceding forced rhythm that they would exactly reproduce it when asked to choose again a convenient rate. (3) Both the forced rhythm and the natural rhythm might exert either equal or differently weighted influences on the rate chosen following the forced rhythm.

Condition I

Table I and Fig. I give the results when 61 subjects were forced to tap at a rate one-half as slow as their own natural rate. The resultant rate lies between the forced slow rate and their original natural rate. It is the third possibility given above that is shown in

[2] The same instructions following the forced rhythms were given in all conditions: "Now tap at a rate that is convenient, natural for you."

TABLE I

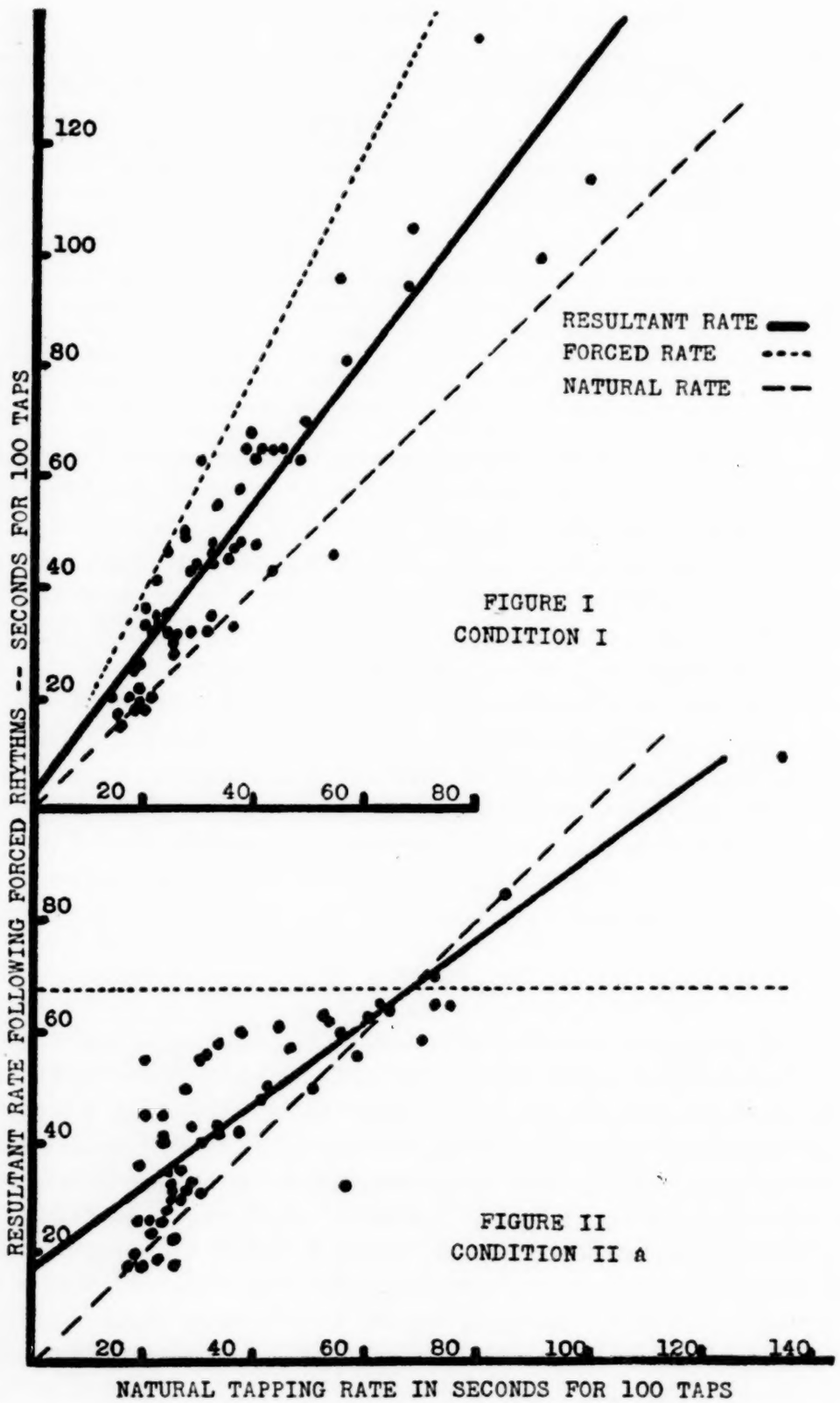
TIME RECORDED IN SECONDS FOR 100 TAPS

N	Natural Rate (3 trials)	Conditions of Forced Rhythms	Resultant Rate after Forced Rhythm			Slope of Regression Lines	
			Mean	C.R.	r	Slope	C.R. of Diff. with Slope of Condition I
61	34.3	I	47.7	8.9	.92	1.33	
55	37.6	II-a	44.9	4.9	.85	.72	6.4
34	39.2	II-b	47.6	3.6	.85	.91	3.5
22	40.4	III	34.8	1.1	.92	.42	10.9

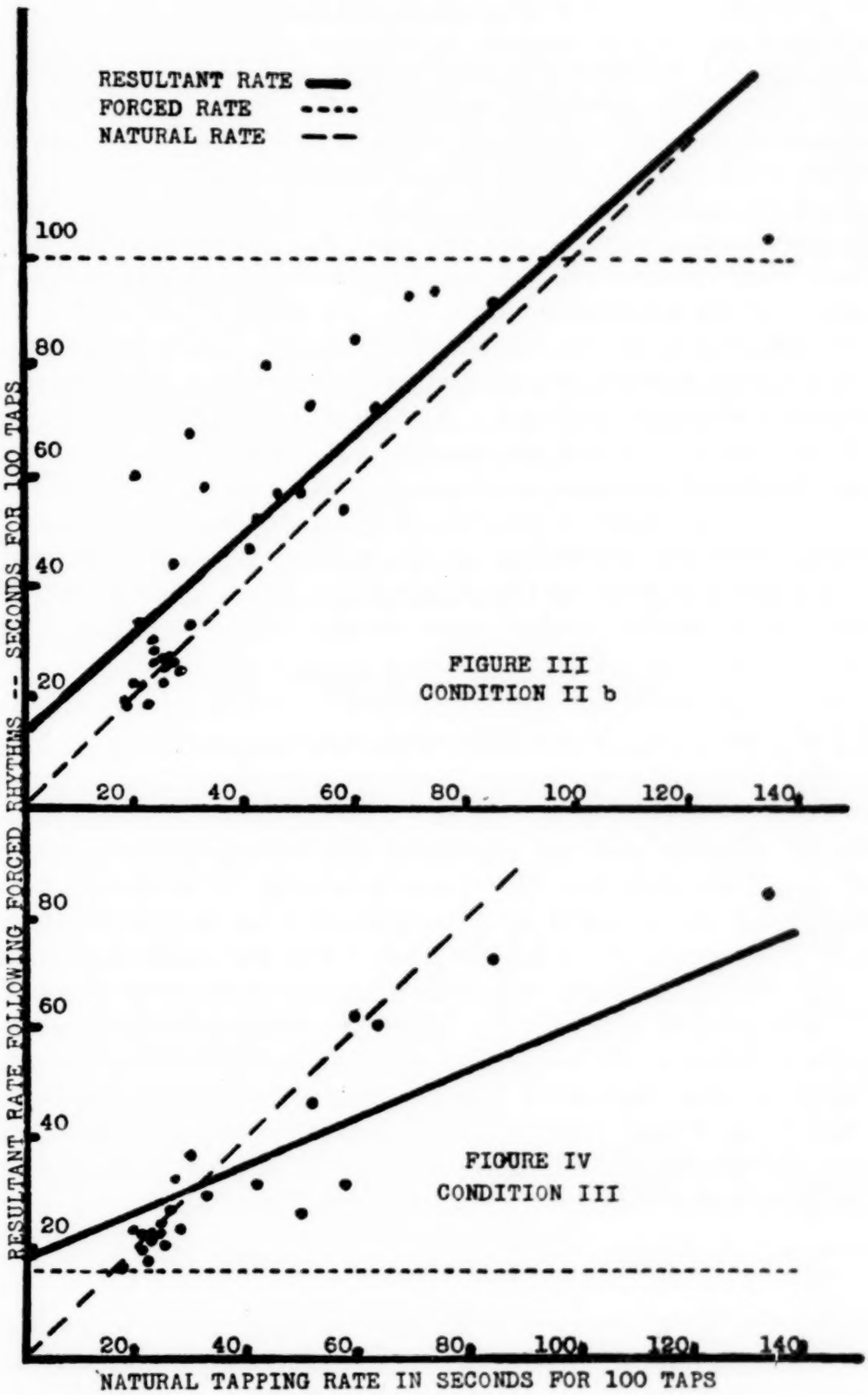
Fig. I. The heavy black line represents the regression line when the resultant rate is predicted from the average score of the three original trials of the natural rate. As can be seen, it lies between the broken line representing the slope if the original natural rate had been reproduced exactly, and the dotted line showing the slope if the forced rhythm had been completely successful in modifying the natural rhythm. A slope of one would have been expected if the highly reliable natural rate had been returned to, and a slope of two if the forced rhythm had been adopted. Instead, the slower rhythm forced on the subjects has so modified their original natural rhythm that the final resultant rate lies between the two with a slope of 1.33. A significant difference is found between the means of the first three trials of the natural rate and the mean of the group when asked to choose again a natural rate following the forced slow rhythm. See Table I.

CONDITION II-a

Fig. II gives the results under Condition II-a. The second group of 55 subjects were uniformly forced to tap to a metronome so set that it took each subject 68 sec. to complete their 100 taps. This particular rate was chosen since it would decrease the rate of the fast subjects, increase the rate of the slow subjects, and not greatly affect those who were near the average original 'natural' rhythm. The result was to swing the regression line around, bring it below the original rhythm at the slow end, and continue it above the fast end. However it does not swing completely around, for it would have then followed the dotted line which indicates the forced rate. The slope is now .72 and significantly different from the preceding slope under Condition I with a critical ratio of 6.4. The forced rhythm in this condition has reliably affected the original natural rhythm, as is also shown by the critical ratio of 4.9 between the means of the resultant rate and the original natural rate.



FIGS. I and II. Relation between natural tapping rate and resultant rate following forced rhythms.



FIGS. III and IV. Relation between natural tapping rate and resultant rate following forced rhythms.

Condition II-b

In Fig. III we have the results from 34 of the 55 subjects used above. As in the preceding condition, we forced a uniform rate on all the subjects, only this forced rhythm was even slower, *i.e.*, all the subjects were asked to tap so that it took them 100 sec. for 100 taps. With all but one of the subjects, this was a slower rate than their preceding natural rhythm. We see that the regression line is lifted slightly from the line which it would have followed if there had been no effect of the intervening activity. Its slope is .91, and is significantly different from the slope of Condition I, again indicating that the new forced rhythm has affected the subjects in a different manner than under the first condition. The critical ratio between the means of the resultant rate and the original natural rate is 3.6. Since the forced rhythm of 100 taps in 100 sec. is even slower than 100 taps in 68 sec., we should expect a greater effect on the resultant rate. The difference between the means of the natural rate and the resultant rate is slightly greater in Condition II-b, which is what we should expect. The smaller critical ratio results from the larger sigma in II-b.

Condition III

In Fig. IV we have the result when the subjects are asked to tap as rapidly as possible for 10 sec. and then asked to choose a rate that was natural, convenient for them. The naturally slow subjects are markedly affected and the regression line swings down toward the rapid rate that they had just been following. The extremely fast subjects were not speeded up proportionally, as they no doubt were already close to the physiological limit. For the group the resultant rate was not significantly different from the average of the natural rate. The critical ratio is 1.1. That the condition has affected the subjects is shown by the critical ratio of 10.9 for the difference between the slope of this regression line and the slope under Condition I. The heavy black line, indicating the resultant rate, is swinging down toward the dotted line which represents the average performance of the subjects when asked to tap as rapidly as possible.

Discussion

These significant changes in the natural tapping rate following the forced rhythms appear to be evidence of induced motor sets affecting the usual adjustment for the natural tapping rate. It should be emphasized that any motor set, apparently set up by a forced rhythm, is quite temporary, although it lasts long enough to affect a highly reliable performance. The significance of the motor sets may lie in the individual differences shown in Figs. I through IV. Particularly in

Figs. II and III can we see a tendency for a curvilinear function, with the naturally fast subjects being least affected by the forced rhythms. The slower individuals are usually markedly affected by the induced rates. In another study recently completed (4), it has been found that individuals who are easily set to continue any rhythm forced on them also tend to perseverate in fields other than motor, *i.e.*, in the sensory and ideational fields. Other studies may reveal more clearly the significance to be attached to the effect of an induced rate upon the natural tapping rate.

Summary

Intervening rates, slower or faster than the natural tapping rate, were forced on subjects who were then asked to return to a rate natural to them. Four conditions were investigated:

I. Effect on the natural tapping rate when subjects were asked to tap at one-half their natural rate.

II-a. Effect on the natural tapping rate when the intervening rate is the same for all subjects, being one-half the average of the preceding group.

II-b. Effect on the natural tapping rate when the intervening rate is the same for all subjects, but in this case only one tap per second.

III. Effect when all the subjects were forced to go faster than their own rate.

The results were as follows:

1. The natural rate, though a highly reliable performance, was in all cases changed in the direction of the forced rhythm. The differences between the means of the original natural rate and the return rate were significant except in Condition III.

2. In no condition do the subjects as a group reproduce the forced rhythm, which shows the consistent effect of the natural rhythm.

3. When the return rate is predicted from the original natural rate, the slope of the regression line varies according to the intervening rate. Significant critical ratios of differences are found between the slopes, using only Condition I as a standard for comparison.

4. Effect of the forced rhythms shows marked individual differences.

(Manuscript received July 7, 1942)

References

1. Allport, G. W., & Vernon, P. E. *Studies in expressive movement.* New York: Macmillan, 1933.
2. Cathcart, P. E., & Dawson, S. Persistence: a characteristic of remembering. *Brit. J. Psychol.*, 1928, **18**, 262–275.
3. Jasper, H. H. Is perseveration a functional unit participating in all behavior processes? *J. soc. Psychol.*, 1931, **2**, 28–51.
4. Rethlingshafer, D. Relationship of tests of persistence to other measures of continuance of activities. *J. abn. soc. Psychol.*, 1942, **37**, 71–82.

A STUDY OF INDIVIDUAL DIFFERENCES IN BINOCULAR COLOR FUSION [1]

BY GLEN GRIMSLEY

University of Southern California

Nearly 140 years have passed since the first important study of binocular color-mixture (1); and, while single investigations have given an impression of conclusiveness, there has been surprisingly little agreement among the numerous investigators (4). If one considers only the question of the binocular-mixture of colored lights, there is quite general agreement as to its possibility, some agreement as to the conditions of fusion, and practically no agreement as to frequency of its occurrence. Considerable is known about the conditions of stimulation (intensity and wave-length separations) which are conducive to fusion, but very little is known about the visual characteristics of those who experience fusion and those who fail to do so. While Rivers (3) suggested that myopic or hypermetropic eyes more easily obtained fusion, and Hecht (2) explained failure to obtain fusion by color-blindness, the relation of peripheral visual defects to failure to experience binocular color fusion has not been established. The purpose of this study was to determine whether or not visual defects (including muscular imbalance and color blindness) were correlated with the inability to experience binocular color-mixture.

Apparatus

The apparatus was similar to that used by Hecht (2), except for the light source. Approximately pure monochromatic lights were obtained by the use of Wratten filters numbers 29 (red), B58 (green), 44A (blue), and 16 (yellow) of the series manufactured by the Eastman Kodak Company for scientific use. The filters were mounted at one end of a light-tight box 3″ × 5″ × 12″ overall, black inside, and the subject looked through the box from the other end. A partition divided the box into two equal lengthwise compartments. Thus, the two eyes were stimulated by differently colored lights.

The filter-box was mounted on a table facing a light diffuser. The light source was an ordinary 75 watt globe [2] placed inside a light-tight box. Light from the globe passed through four

[1] The experimental results of this study were contained in a thesis presented by the writer to the Faculty of the Department of Psychology, The University of Southern California, in partial fulfillment of the requirements for the degree of Master of Arts, May 1942. The writer wishes to express his appreciation to Doctor Neil Warren for his help and for his keen interest in the problem. The study was made possible by the full coöperation of Doctor Ernest Hutchinson and members of the staff of the Los Angeles School of Optometry. The writer is especially indebted to Doctor Leo Davidson for his assistance in interpreting the results of the optometric examinations and to Mr. Gregory Poulogianis for his suggestions regarding apparatus.

[2] Following the advise of the Eastman Laboratories, a day-light globe was used with the red and green, and an ordinary frosted globe with the blue and yellow filters.

single-strength panes of opal glass at intervals of one inch. The 'outside' pane, marked with a black fixation point in its center, faced the filter-box. The entire diffuser was movable, so that the distance from the subject's eyes to the fixation point on the illuminated window could be varied to suit the individual. In order to keep the stimulating lights as nearly monochromatic as possible, the apparatus was set up in a dark-room.

When a subject looked at the fixation point through the filter-box he saw a rectangular portion of the illuminated window which appeared to be divided into three colored sections as described by Hecht (2). The two ends of the rectangle corresponded to the filter colors and, in cases experiencing fusion, the center section was a mixture of the two—yellow with the red and green filters, and white with the blue and yellow.

Experimental Procedure

This study involved two examinations of each subject: (1) a regular optometric examination which was given by the Los Angeles School of Optometry, and (2) an investigation, by means of the apparatus described above, of each subject's ability to fuse binocularly red and green, and blue and yellow lights. Fifty-six subjects (fifty normal, five red-green blind, and one red-green and blue blind [3]) were used.

Each subject was given the Ishihara color-blindness test. He was then brought into the dimly-lighted dark-room and given instructions while he was becoming adapted to the reduced illumination. The subject was shown a black and white outline-drawing of the three rectangular sections of color that he would be able to see through the filter-box. He was then seated at the end of the table and the experimenter said: "Assume a comfortable position, rest your arms on the table if you wish. Look through the filter-box and fixate the black spot on the surface of the illuminated window. Tell me when you see the three rectangular sections that I have shown you." When the subject had become accustomed to the apparatus and had no difficulty in seeing the three sections of color, the room was darkened except for the light coming through the opal-glass window. The subject was further instructed: "Tell me what color the center section is. Does it change? Report the changes. Is it at any one time a solid color?" This procedure was carried through with both sets of filters, under normal visual conditions and with vision confined to the foveal area by means of the 'pin-holes' used by optometrists.

It was considered necessary to check two factors that might have affected the results. These were (1) eye dominance, and (2) the correction of visual defects by glasses. Hence, in every case of failure to obtain fusion, the filters were reversed. In no case did this lead to fusion. Eighteen subjects wore glasses; however, there were no significant fusion differences between the group wearing and the group not wearing glasses. Those wearing glasses did equally well without their glasses, except for one subject who 'saw double' without his glasses.

Results

Of the 50 normal subjects, 66 percent reported 'yellow' to red and green binocular stimulation, and 38 percent reported 'white' to blue and yellow binocular stimulation. With stimulation confined to the foveal area by means of 'pin-holes,' these percentages were increased to 84 and 90 respectively. With no limit on the area of the retina stimulated, 30 percent of the normal subjects reported a binocular mixture of both red and green, and blue and yellow. With stimulation confined to the foveal area, this percentage was increased to 80.

Contrary to Hecht's findings (2), all 6 red-green blind subjects experienced fusion of red and green. Four (67 percent) called both

[3] Two of these *S*s failed all of the Ishihara charts, two named one each correctly, and the other two responded correctly to four of the thirteen charts. The *S* designated as being also blue blind could not distinguish between blues and grays, and persistently spoke of the blue light, used in the fusion test, as gray.

TABLE I

FREQUENCY OF OCCURRENCE OF BINOCULAR FUSION IN NORMAL AND IN FOVEAL VISION (N = 50)

Subjects Experiencing Fusion of	Red and Green		Blue and Yellow	
	Normal Vision	Foveal Vision	Normal Vision	Foveal Vision
Number	33	42	19	45
Percent	66	84	38	90

red and green 'yellow.' Four (67 percent) reported 'white' to blue and yellow binocular stimulation. The red-green and blue blind individual failed to obtain a fusion of blue and yellow. He reported that the center section was divided vertically—one half yellow and the other gray.[4]

Of the 50 normal subjects, 15 (group A) achieved fusion of both pairs of colors under normal conditions[5] and 13 (group B) failed to obtain fusion of either pair. Group A showed slightly less corneal astigmatism than group B. The mean difference in amount of corneal astigmatism in diopters for the two eyes was .2670 for group A, and .4519 for group B. The difference between these means is not significant (C.R. = 1.67). Groups A and B were further compared in regard to a muscular fatigue factor as measured by the ratio of the 'make' score to the 'break' score (both scores in prism diopters) on the tests of convergence and abduction. Group A showed less fatigue (or a greater reserve of muscular energy) in the muscles responsible for convergence than did group B. The difference between the means of the two groups was .167, the σ_D was .0904. The critical

TABLE II

COMPARISON OF VARIOUS OPTOMETRIC TEST RESULTS FOR GROUPS A * AND B *

	Group A* N = 15	Group B* N = 13	Diff.	σD	C.R.
Difference in the two eyes in amount of corneal astigmatism	.2670	.4519	.1849	.1106	1.67
Amount of convergence fatigue †	.497	.330	.167	.0904	1.85
Amount of abduction fatigue †	.513	.423	.090	.0756	1.19
Total amount of fatigue † convergence or abduction	.423	.262	.161	.0650	2.48

* Group A—Those who experienced fusion of both pairs of colors under normal conditions. Group B—Those who experienced fusion of neither pair of colors under normal conditions.

† Low scores indicate greater fatigue.

[4] Vertical divisions of the center section were accompanied by side-to-side reversals of the colors in this section.

[5] No limit on the area of the retina stimulated.

ratio (1.85) is not significant. In the abduction test the difference was again in favor of group A, and again the critical ratio (1.19) is not significant. However, it was thought that a fatigued condition either of the muscles responsible for convergence or of those responsible for abduction would probably lead to fixation difficulties and increased likelihood of stimulation of non-corresponding points of the two retinas. For this reason the results of the two tests were combined. Here the difference between the means for groups A and B is .161, the σ_D is .0650 and the critical ratio of 2.48 is reasonably significant. Group A shows significantly less fatigue of the muscles responsible for convergence and abduction.

Discussion

While this study was only exploratory, the findings make it seem probable that many failures to experience binocular color fusion are due to the inability to maintain continuous stimulation of corresponding points on the two retinae, because of imbalance or fatigue of the rectus and obliquus muscles. Insofar as refractory defects are responsible for failure to experience fusion, they seem to be of a kind which are not corrected by glasses. However, 'maximum correction' by the use of 'pin-holes' resulted in approximately doubling the frequency of fusion, and greatly reduced retinal rivalry. All cases reporting failure to experience binocular color-mixture in foveal vision reported either a division of the visual field or the wrong color—gray, white, and orange instead of yellow, and gray and pale yellow instead of white. Thus, it would seem that while a small number of the failures was associated with peculiarities of color reception, the great majority were associated with defects of the accessory sense-apparatus. This view is further supported by the apparent association of the ability to fuse red and green, and the ability to fuse blue and yellow.[6]

In spite of the fact that to regard yellow as a central response removes the objection to the 'three component' part of the Young-Helmholtz theory (2), the reports of the color-blind subjects support the Ladd-Franklin theory. All of the red-green blind subjects experienced binocular fusion, two-thirds of them called both red and green 'yellow,' and the blue-blind subject called blue 'gray,' which is exactly what one would expect according to the evolutionary theory.

[6] The probability of dependence equals approximately .85 by the chi-square test. *Cf.* Garrett, H. E. *Statistics in psychology and education.* (2nd ed.) New York: Longmans, Green, 1938, p. 377.

Summary

1. Fifty normal and 6 color-blind subjects were tested in the ability to fuse binocularly red and green, and blue and yellow monochromatic lights by having them look at an illuminated opal glass window through a box, painted black inside and bearing a Wratten filter in front of each eye.

2. Of the normals, 66 percent reported 'yellow' to red and green binocular stimulation under normal conditions, and 38 percent reported 'white' to blue and yellow binocular stimulations. By confining stimulation to the foveal area these percentages were increased to 84 and 90 respectively. With no limit on the areas of the retinas stimulated, 30 percent of the normal subjects reported achieving a binocular mixture of both red and green, and blue and yellow. By confining the stimulation to foveal area, this percentage was increased to 80.

3. The 50 subjects were given optometric examinations by the Los Angeles School of Optometry. The results of certain parts of these examinations for the 15 subjects who achieved fusion of both pairs of colors under normal conditions were compared with the results of the 13 subjects who failed to obtain fusion of either pair. The two eyes of the successful subjects were found to be more nearly alike than those of the unsuccessful subjects in regard to amount of corneal astigmatism. However, the critical ratio (1.67) was not significant. Comparisons of the two groups in regard to convergence and abduction fatigue also showed differences in favor of the successful group, with critical ratios of 1.85 and 1.19 respectively. When the fatigue factor was treated as a whole (both convergence and abduction) the differences became reasonably significant (C.R. = 2.48). It should be noted that, while only the last critical ratio is significant, the differences on all the tests are consistently in favor of the group which obtained fusion. The probability of such an occurrence by chance is extremely remote.

4. The 6 red-green blind subjects were given the tests for binocular fusion only. All 6 of them reported 'yellow' to red and green binocular stimulation, and 4 of them reported 'white' to blue and yellow stimulation. Thus, 100 percent of the red-green blind individuals experienced a fusion of red and green, and 67 percent of blue and yellow. The corresponding percentages of the normal group were 66 and 38 respectively. Four (67 percent) of the red-green blind individuals saw both red and green as 'yellow.'

5. It is suggested that, while a small percent of failures to experience binocular color fusion may result from peculiarities of the color receptors, the largest percent of failures may result from defects in

the accessory sense-apparatus which make impossible the continuous stimulation of corresponding points on the two retinae.

(Manuscript received July 6, 1942)

References

1. Haldat, C. N. A. Sur la double vision. *J. de physique*, 1806, **63**, 387. (Cited by S. Dawson, *Brit. J. Psychol.*, 1915–1917, **8**, 510.)
2. Hecht, S. On the binocular fusion of colors and its relation to the theories of color vision. *Proc. nat. Acad. Sci.*, 1928, **14**, 237–241.
3. Rivers, W. H. R. On binocular color mixture. *Proc. Cambridge phil. Soc.*, 1895, **8**, 273–277.
4. Woodworth, R. S. *Experimental psychology*. New York: Henry Holt and Co., 1938, pp. 574–575.

CONTINUOUS RECORDING OF PULSE AND BLOOD PRESSURE

BY EDWARD GIRDEN[1]

Brooklyn College *

In the course of studying generalized somatic conditioned responses in animals subjected to erythroidine (**2**), it became necessary to procure continuous records of blood pressure and pulse rate. For such experiments, which require long continuous recording, the smoked-paper kymograph is impractical and a polygraph is indicated. The Parker polygraph [2] has been entirely adequate and may be used in either the horizontal or, as illustrated in Fig. 1, the vertical position. It has 3 speeds, 7.5 cm, 30 cm and 45 cm per min, respectively, and takes a 300-ft roll of recording paper 6 in. wide.

The success in obtaining continuous polygraph recordings of delicate responses depends upon avoiding undue inertia. This was accomplished by using wax paper with hot-point needles constructed from thin nichrome-wire. In the present experiments, blood pressure and pulse were taken in the usual way from the carotid artery by means of a mercury manometer (*cf.* **3**). Hot-points were attached to it and the signal magnets. The leads, being of thin wire, did not produce undue weight or inertia and, particularly in the case of the mercury-manometer hot-point, provided sufficient tension so as to maintain proper contact. Some preliminary adjustments might be required but, once set, the apparatus would then function without need for for further attention for the duration of the experiment. Sudden large changes in blood pressure were quickly compensated for, since the manometer was held by a rack and pinion which could be adjusted in height readily without disturbing the set-up.

Details of the apparatus were as follows:

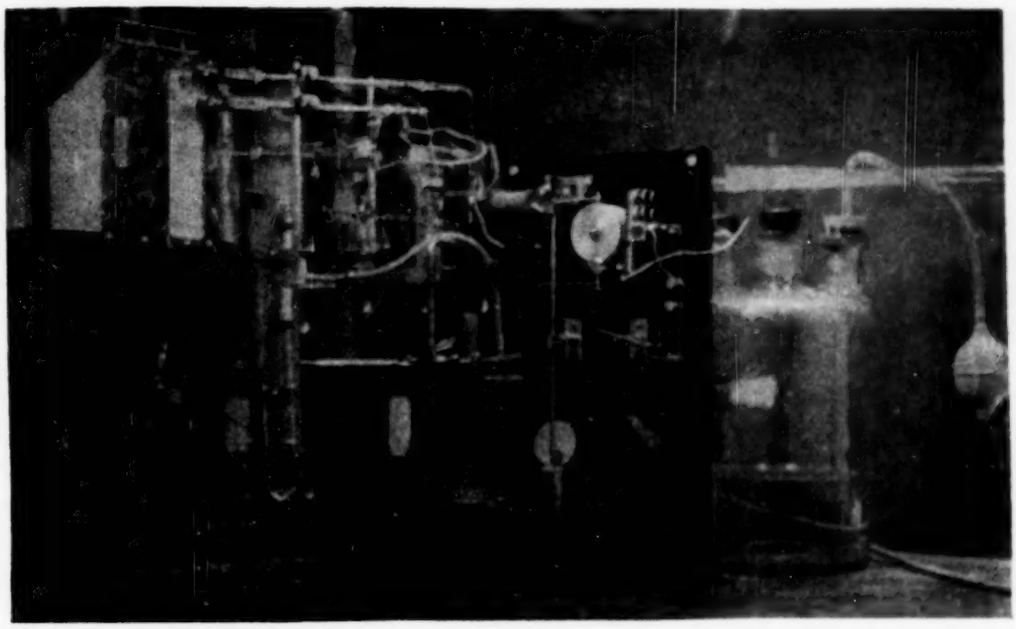
Hot-point electrodes were constructed from thin nichrome-wire and soldered respectively to the two vibrators of the 10 mm Harvard signal magnets and to the tip of the cross-arm of the upright resting in the open side of mercury manometer. The signal magnets, clamped to the metal support on the polygraph, formed common ground for

[1] Fellow of the John Simon Guggenheim Memorial Foundation, 1941–42. Special aid from the Penrose Fund of the American Philosophical Society and the Elizabeth Thompson Science Fund is gratefully acknowledged.

* Communication No. 49 from the Physiological Psychology Laboratory, University of Rochester, maintained by aid of the Research Council, American Otological Society.

[2] O. Parker Engineering Products, 16 West 22nd St., New York City.

$1\frac{1}{2}$ volt dry-cells (B_1, B_2). Thin copper-wire leads, 24 gauge with cotton-thread insulation, were soldered to the free ends of the hot-points and served as active electrodes. The amount of heat produced was controlled by changing the position of the 'active' lead



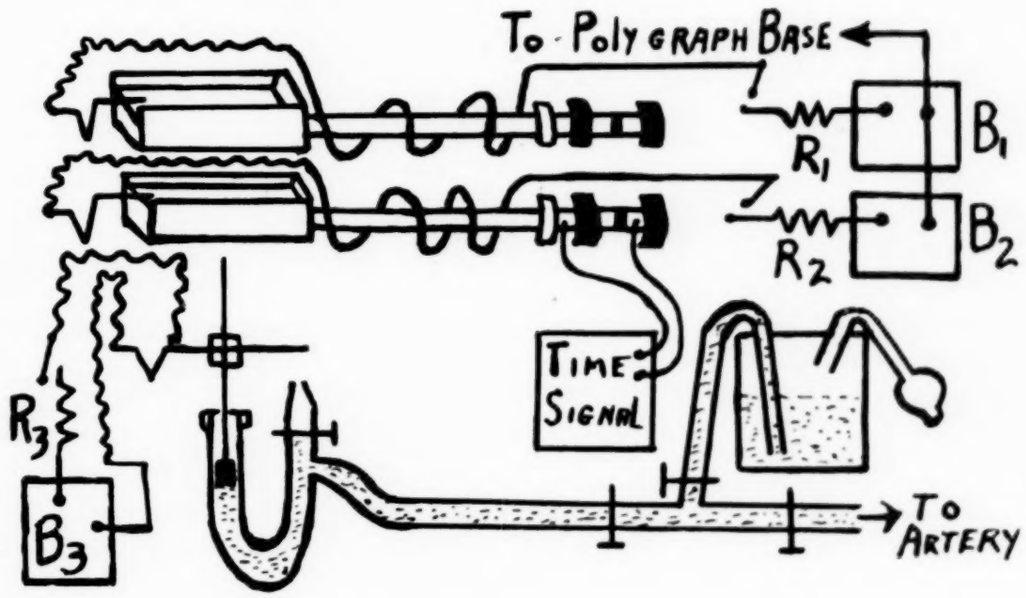


FIG. 1. Apparatus for continuous recording of pulse and blood pressure.

(which was attached to a small battery clip) on the resistances, R_1 and R_2, made up of 4 in. lengths of nichrome wire.

The mercury manometer was independent of the polygraph and supported by a Gaertner rack-and-pinon. This hot-point required two leads to its battery, B_3, and the heat was controlled by R_3.

Long leads produced sufficient tension to keep this hot-point in contact with the wax paper without impeding action of the plunger resting on the mercury column.

The time signal was produced by a Harvard Clock (battery source not indicated). Similar leads from the stimulus, *e.g.*, tone or light, were connected to the binding posts of the second signal magnet. In actual practice, with even three signal magnets, it was possible to record, without difficulty, variations in blood pressure up to 150 mm.

In actual experimental work, records have been taken without interruption for as long as six hours. When necessary, a new roll of paper can be inserted with little loss of time. We have found the recording speed of 7.5 cm per min to be a little too slow to record the pulse beat (*cf.* **3**); with medium speed, 30 cm per min, the pulse appears very well (*cf.* **4**).

This technique also has proven advantageous in recording the pupillary response in human subjects by means of a pupillometer. In previous studies (*cf.* **1**), a 'cold' mechanical recorder was employed which made its impressions on the wax-paper by pressure-contact. With the use of the hot-point, requiring almost imperceptible contact, undue weight and inertia are avoided with the result that distortion is at a minimum. The results of this study are to be reported in the near future. Other uses for this technique come to mind readily. The reliability of this method has been proven in actual experimental research and it should have a variety of useful applications.

(Manuscript received June 22, 1942)

References

1. Baker, L. E. The pupillary response conditioned to subliminal auditory stimuli. *Psychol. Monogr.*, 1938, **50**, No. 223, pp. iii + 32.
2. Girden, E. Generalized conditioned responses under curare and erythroidine. *J. exp. Psychol.*, 1942, **31**, 105–119.
3. Girden, E. The dissociation of blood pressure conditioned responses under erythroidine. *J. exp. Psychol.*, 1942, **31**, 219–231.
4. Girden, E. Role of the response mechanism in learning and 'excited emotion.' *Amer. J. Psychol.*, 1943, **56**, No. 1. (In press.)

DISCUSSION

DR. SKINNER ON THE 'DISINTEGRATION OF PERCEPTUAL PRINCIPLES' OF PSYCHOTICS

BY G. K. YACORZYNSKI

Northwestern University

In the Zenith experiment one of two objects (circle and cross, heads and tails, etc.) was selected by chance and the radio listeners were asked to guess which of the two objects had been chosen. Five trials were used for each of the two alternatives. Recently I have described the results of an experiment simulating this method on manic-depressive and schizophrenic patients (4). I found that my controls selected patterns similar to that of the Zenith population reported by Goodfellow (2), but that the psychotics differed from the controls, although the responses of the manic-depressives and the schizophrenics were similar.

In a recent discussion of these data, Dr. Skinner (3) attributes to me the ambiguous statement of 'disintegration of perceptual principles' as my description of the differences which were found between the controls and the psychotics. Since this statement as worded by Skinner does not appear in my paper, and is not conveyed by the context of the discussion, it is necessary to show the manner in which Skinner's discussion and mine differ.

Skinner concludes that the results of the psychotics can be described by a single statement that of a 'permanent tendency to alternate.' My statement reads, "The most preferred pattern of the psychotics was the 12121 sequence." No discrepancy exists between these statements.

As to the conditions which produce these responses Skinner states, "In making a series of five calls between the two alternatives the psychotic differs from the normal in being uninfluenced by his past behavior." My concluding statement reads, "It is suggested that the psychotics rather than modifying their responses to conform to the special conditions demanded by the instructions resort to more basic patterns which are used by the normal in the absence of any inhibitory conditions." Although a great deal of similarity exists between Skinner's conclusion and mine, in the fact that both statements agree that the psychotics remain uninfluenced by external conditions, certain differences are present. To clarify these differences it is necessary to understand the nature of the responses of the normal controls.

Skinner finds that the preceding pair of calls (with some exceptions) influences the next response in the normal individual. For example, if in the five trials two calls have been similar, there is a tendency to alternate the next call, or if the calls have been alternated, then the opposite condition holds. According to this viewpoint the antecedent conditions (disregarding the first call) which produce the responses of the normal are the calls which he has already made.

Skinner, however, disregards the fact that the individual is actually attempting to guess the manner in which chance may determine which one of the two alternatives have been selected. It is this experimental condition which determines the nature of the responses of the individual, since if he were not confronted with the necessity of guessing the outcome of the chance selection of one of two alternatives he would show an entirely different sequence of responses. It is necessary to have these or similar experimental conditions in order that the similarity or dissimilarity of two calls should influence the future call in the manner already described.

The incomplete statement of the antecedent conditions which determine the responses in the selection of alternatives in the normal leads Skinner to conclude that the psychotic fails to be influenced by the preceding calls. It is more nearly correct, however, to state as I have done that the psychotic fails to be influenced by the 'special conditions demanded by the instructions.' The psychotic is not influenced by the preceding calls because he fails to respond to the whole experimental situation.

If the psychotic is uninfluenced by the conditions of the experiment, then another question presents itself. Why does he select a certain predictable sequence, or in other words, what are the determiners of the tendency to alternate his responses? If there were no antecedent or concomitant conditions influencing the response then no definite pattern would emerge. The work of Fernberger (1) indicates where the solution to this question may be found, in the fact that normal individuals will show a tendency to alternate their responses if they are not required to predict the alternatives which will occur from chance selection. It is for this reason that I make the statement that the psychotics "resort to the more basic patterns which are used by the normal in the absence of any inhibitory conditions." This question does not present itself to Skinner because he is satisfied merely to describe the relationships which exist between the separate calls made in the guessing experiment, and not the conditions which produce the sequence of responses.

Thus, Skinner's contribution is the analysis of the manner in which, given a definite experimental situation, each choice in the selection of alternatives influences the subsequent response. He assumes that the preceding responses are the sole determiners of what the individual's future response will be. This is a perfectly acceptable statement if one is satisfied to disregard all of the important antecedents determining the response. If these conditions are considered then it is evident that the whole experimental situation has determined the sequence of responses. Skinner's objection to the use of the term 'pattern' in the sense of 'a single organized act,' in referring to the sequence of responses which the individual makes, follows as a consequence of his interest in analyzing the pattern of guesses into discrete events, with a disregard of the important determiners of the total pattern.

(Manuscript received October 15, 1942)

References

1. Fernberger, S. W. Interdependence of judgments within the series for the method of constant stimuli. *J. exp. Psychol.*, 1920, 3, 126–150.

2. GOODFELLOW, L. D. A psychological interpretation of the results of the Zenith radio experiments in telepathy. *J. exp. Psychol.*, 1938, **23**, 601–632.
3. SKINNER, B. F. The processes involved in the repeated guessing of alternatives. *J. exp. Psychol.*, 1942, **30**, 495–503.
4. YACORZYNSKI, G. K. Perceptual principles involved in the disintegration of a configuration formed in predicting the occurrence of patterns selected by chance. *J. exp. Psychol.*, 1941, **29**, 401–406.

REPLY TO DR. YACORZYNSKI

BY B. F. SKINNER

University of Minnesota

Dr. Yacorzynski is justified in asking me to remove the quotation marks from 'disintegration of perceptual principles,' but I cannot agree that I have misrepresented him or that the notion is 'not conveyed by the context' of his discussion. He speaks of the "manner in which the normal perceptual process functions and disintegrates"; and if I have interchanged *process* and *principle*, it is only in following Dr. Yacorzynski's example. He speaks of "the principles . . . which are affected" (p. 401) and the "process which is affected" (p. 402) in referring to the same phenomenon; and after mentioning the possibility that the "normal preference for non-symmetry has simply disintegrated," he speaks of non-symmetry as a 'principle.'

Nor can I agree with his claim that there is a 'great deal of similarity' in our conclusions. It is true that "no discrepancy exists" between my statement that there is a permanent tendency to alternate and Yacorzynski's that "the most preferred pattern of the psychotics was the 12121 sequence." But they are by no means the same statement. There is no discrepancy because the latter is one consequence of the former. There are fifteen other consequences, to which Yacorzynski's statement does not refer. My statement that the psychotic remains 'uninfluenced by his past behavior' is certainly not the same as Yacorzynski's that psychotics fail to 'modify their responses to conform to the special conditions demanded by the instructions,' since one refers to stimulation generated by the subject and the other to an external source.

Far from disregarding the fact that the individual is "actually attempting to guess the manner in which chance may determine which one of the two alternatives has been selected," my paper is concerned with analyzing exactly that case. My second paragraph contains in effect a definition of guessing. Again, I do not assume "that the preceding responses are the sole determiners of what . . . future responses will be." I note the possibility of 'abiding preferences,' 'biased preliminary conditions,' and "trivial circumstances which cancel out in the long run and are spoken of as 'chance.'" A second guess is said to be under the *additional* control of the first, and so on. As it turns out, this control is all-important; although I do not assume that it is the sole determiner, it is practically shown to be so.

Dr. Yacorzynski prefers to deal with the 'total situation' and with a series of five guesses as a 'configuration.' My own preference is for a more

detailed analysis of both stimulus and response. I specifically admit the possibility of analysis at other levels (p. 499) and am only concerned with testing the descriptive power of the level I prefer. As Dr. Yacorzynski says, "Skinner's objection to the use of the term pattern in the sense of 'a single organized act' in referring to the sequence of responses which the individual makes follows as a consequence of his interest in analyzing the pattern of guesses into discrete events." I am satisfied with that statement of our relative positions and have no wish to dispute such a matter. However, he adds: "with a disregard of the important determiners of the total pattern." This tiresome complaint has seldom been made under less auspicious circumstances. It is often difficult to answer the charge that analysis 'ignores' something or other, because in our present state of knowledge it is difficult to prove that everything has in fact been accounted for. But here the shoe is on the other foot. Dr. Yacorzynski has not challenged my contention that the prediction of guessing in terms of the history of alternation is more exact than his own prediction in terms of the total situation. Just how can the ignoring of any relevant matter lead to a more adequate treatment of the facts?

(Manuscript received October 15, 1942)